



SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Goodman, Corey S.
Kidd, Thomas
Mitchell, Kevin
Tear, Guy
- (ii) TITLE OF INVENTION: Robo: A Novel Family of Polypeptide and Nucleic Acids
- (iii) NUMBER OF SEQUENCES: 13
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
 - (B) STREET: 75 DENISE DRIVE
 - (C) CITY: HILLSBOROUGH
 - (D) STATE: CALIFORNIA
 - (E) COUNTRY: USA
 - (F) ZIP: 94010
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: OSMAN, RICHARD A
 - (B) REGISTRATION NUMBER: 36,627
 - (C) REFERENCE/DOCKET NUMBER: B98-006
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (650) 343-4341
 - (B) TELEFAX: (650) 343-4342

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4188 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGCATCCCA	TGCATCCCGA	AAACCACGCC	ATCGCCCGGA	GCACGAGCAC	CACTAATAAC	60
CCATCTCGCA	GTCGGAGCAG	CAGGATGTGG	CTCCTGCCCC	CCTGGCTGCT	CCTCGTCCTG	120
GTGGCCAGCA	ATGGCCTGCC	AGCAGTCAGA	GGCCAGTACC	AATCGCCACG	TATCATCGAG	180
CATCCACCGG	ATCTGGTCGT	TAAGAAGAAT	GAACCCGCCA	CGCTCAACTG	CAAAGTGGAG	240
GGCAAGCCGG	AACCCACCAT	TGAGTGGTTT	AAGGATGGCG	AACCCGTCAG	CACCAACGAA	300
AAGAAATCGC	ACCGCGTCCA	GTTCAAGGAC	GGCGCCCTCT	TCTTTTACAG	GACAATGCAA	360
GGCAAGAAGG	AGCAGGACGG	CGGAGAGTAC	TGGTGCGTGG	CCAAGAACCG	AGTGGGCCAG	420
GCCGTTAGTC	GCCATGCCTC	CCTCCAGATA	GCTGTTTTGC	GCGACGATTT	TCGCGTGGAG	480
CCCAAAGACA	CGCGAGTGGC	CAAAGGCGAG	ACGGCTCTGC	TGGAGTGTGG	GCCGCCCAAA	540
GGCATTCCAG	AGCCAACGCT	GATTTGGATA	AAGGACGGCG	TTCCCTTGGA	CGACCTGAAA	600
GCCATGTCGT	TTGGCGCCAG	CTCCCGCGTT	CGAATTGTGG	ACGGTGGCAA	CCTGCTGATC	660
AGCAATGTGG	AGCCCATTTGA	TGAGGGCAAC	TACAAGTGCA	TTGCCAGAA	TCTGGTAGGC	720
ACCCGCGAGA	GCAGCTATGC	CAAGCTGATT	GTCCAGGTCA	AACCATACTT	TATGAAGGAG	780

CCCAAGGATC	AGGTGATGCT	CTACGGCCAG	ACAGCCACTT	TCCACTGCTC	AGTGGGCGGT	840
GATCCGCCGC	CGAAAGTGTT	GTGGAAAAAG	GAGGAGGGCA	ATATTCCGGT	GTCCAGAGCG	900
CGAATCCTTC	ACGACGAGAA	AAGTTTAGAG	ATATCCAACA	TAACGCCCAC	CGATGAGGGC	960
ACCTATGTCT	GCGAGGCACA	CAACAATGTC	GGTCAGATCA	GCGCTAGGGC	TTCTCTTATA	1020
GTCCACGCTC	CGCCGAACCT	TACGAAAAGA	CCCAGTAACA	AGAAAGTGGG	ACTAAATGGG	1080
GTTGTCCAAC	TACCTTGTCAT	GGCCTCCGGA	AACCCTCCGC	CGTCTGTATT	CTGGACCAAG	1140
GAAGGAGTAT	CCACTCTTAT	GTTCCCAAAT	AGTTCGCACG	GAAGGCAGTA	TGTGGCTGCC	1200
GATGGAACCT	TGCAGATTAC	GGATGTGCGG	CAGGAAGACG	AAGGCTACTA	TGTGTGTTCC	1260
GCTTTCAGTG	TAGTCGATT	CTCTACAGTA	CGGGTTTTCC	TGCAAGTCAG	CTCGGTAGAC	1320
GAGCGTCCAC	CTCCGATTAT	TCAAATCGGA	CCTGCCAATC	AAACACTGCC	CAAGGGATCA	1380
GTTGCTACTT	TACCCTGTCT	GGCCACTGGA	AATCCCAGTC	CCCGTATCAA	GTGGTTCCAC	1440
GATGGACATG	CCGTACAAGC	GGGCAATCGA	TACAGCATCA	TCCAAGGAAG	CTCACTGAGA	1500
GTCGATGACC	TTCAACTAAG	TGACTCTGGT	ACCTACACCT	GCACTGCATC	TGGCGAACGA	1560
GGAGAAACTT	CCTGGGCTGC	CACACTAACG	GTGGAAAAAC	CCGGTTCTAC	ATCTCTTCAC	1620
CGGGCAGCTG	ATCCTAGCAC	TTATCCTGCT	CCTCCAGGAA	CACCTAAAGT	CCTGAATGTC	1680
AGTCGCACCA	GCATTAGTCT	TCGTTGGGCT	AAAAGCCAAG	AGAAACCCGG	AGCTGTGGGC	1740
CCAATCATTG	GATACACTGT	AGAGTACTTC	AGTCCGGATC	TGCAAACTGG	TTGGATTGTG	1800
GCTGCCCATC	GAGTCGGCGA	CACTCAAAGTC	ACTATCTCGG	GTCTCACTCC	TGGCACTTCG	1860
TATGTGTTCC	TAGTTAGAGC	TGAGAATACT	CAGGGTATTT	CTGTGCCTTC	CGGCTTATCA	1920
AATGTTATTA	AAACCATTGA	GGCAGATTTT	GATGCAGCTT	CTGCCAATGA	TTTGTGAGCA	1980
GCTCGAACTT	TGCTGACAGG	AAAGTCGGTG	GAGCTAATAG	ATGCCTCGGC	TATCAATGCT	2040
AGTGCCGTTA	GACTTGAGTG	GATGCTCCAC	GTGAGCGCTG	ATGAGAAATA	CGTAGAGGGC	2100
CTGCGCATAC	ACTATAAGGA	TGCCAGTGTA	CCATCCGCAC	AGTATCACTC	GATCACTGTT	2160
ATGGATGCCT	CTGCAGAATC	GTTTGTGGTG	GGAAACCTTA	AGAAGTACAC	CAAGTATGAG	2220
TTCTTCCCTAA	CACCCTTTTT	TGAGACAATT	GAAGGACAGC	CCAGTAACTC	CAAGACAGCC	2280
CTCACCTATG	AAGATGTTCC	CTCCGCACCA	CCGGATAACA	TTCAGATTGG	CATGTACAAC	2340
CAAACAGCCG	GTTGGGTGCG	TTGGACTCCG	CCACCCTCCC	AGCACCACAA	TGGCAATTTG	2400
TATGGCTACA	AGATTGAGGT	CAGCGCCGGT	AACACCATGA	AGGTGCTGGC	CAATATGACT	2460
CTTAATGCTA	CCACCACATC	TGTGCTCCTA	AATAACCTAA	CCACCGGAGC	TGTGTACAGC	2520
GTGAGGTTGA	ACTCCTTTAC	CAAGGCAGGA	GATGGACCTT	ACTCCAAACC	GATATCACTA	2580
TTCATGGACC	CCACCCATCA	TGTGCATCCG	CCACGGGCAC	ATCCAAGCGG	CACCCATGAT	2640
GGGCGACATG	AGGGACAGGA	TCTCACGTAT	CATAACAATG	GCAACATACC	ACCTGGCGAC	2700
ATTAATCCCA	CCACTCATAA	AAAGACCACT	GACTACCTAT	CTGGACCGTG	GCTAATGGTG	2760
CTGGTCTGCA	TCGTTCTTCT	AGTCCTGGTT	ATTTCCGGCG	CTATTTTCGAT	GGTCTACTTC	2820
AAGCGCAAGC	ATCAAATGAC	CAAGGAATTG	GGTCACTTAA	GTGTGGTCAG	TGACAACGAA	2880
ATAACCGCAT	TAAATATCAA	TAGCAAAGAG	AGCCTTTGGA	TAGACCATCA	TCGTGGATGG	2940
CGAACTGCCG	ATACTGACAA	AGACTCAGGA	TTAAGCGAAT	CGAAGCTACT	ATCCACGTT	3000
AACAGCAGTC	AATCCAACCTA	CAATAACTCC	GATGGAGGAA	CCGATTATGC	AGAAGTTGAC	3060
ACCCGTAACC	TTACCACCTT	CTACAATTGT	CGCAAGAGCC	CCGATAATCC	CACGCCGTAC	3120
GCCACCACTA	TGATCATTGG	TACCTCTTCC	AGTGAGACCT	GCACCAAGAC	AACATCTATA	3180
AGTGCCGATA	AGGACTCGGG	AACTCATTCG	CCCTATTCTG	ACGCATTTGC	CGGTGAGGTG	3240
CCAGCGGTTT	CTGTTGTCAA	ATCCAACCTAT	CTTCAGTATC	CGGTTGAACC	GATCAACTGG	3300
TCAGAGTTTC	TACCCCGGCC	GCCAGAACAC	CCACCTCCGT	CTTCTACCTA	TGGATACGCA	3360
CAAGGATCTC	CTGAATCTTC	GCGGAAGAGC	TCCAAAAGCG	CAGGTTCCGG	CATTTCTACA	3420
AATCAAAGCA	TTCTGAACGC	ATCCATACAC	AGCAGCTCCT	CGGGCGGCTT	TTCAGCTTGG	3480
GGAGTATCGC	CCCAATATGC	TGTCGCCTGT	CCACCGGAAA	ACGTTTATAG	CAATCCGCTG	3540
TCGGCAGTGG	CTGGCGGCAC	CCAGAACCGC	TATCAGATAA	CGCCCACAAA	CCAACATCCG	3600
CCACAGTTAC	CGGCCTACTT	TGCCACCACG	GGTCCAGGAG	GAGCTGTACC	ACCCAACCAC	3660
CTGCCATTTG	CCACACAGCG	TCATGCAGCC	AGCGAGTACC	AGGCTGGACT	GAATGCAGCG	3720
CGATGTGCCC	AAAGCCGCGC	CTGCAACAGC	TGCGATGCCT	TGGCCACACC	CTCGCCCATG	3780
CAACCCCCAC	CGCCAGTTCC	CGTACCCGAG	GGCTGGTACC	AACCGGTGCA	TCCCAATAGC	3840
CACCCGATGC	ACCCGACCTC	CTCCAACCAC	CAGATCTACC	AGTGCTCCTC	CGAGTGCTCG	3900
GATCACTCGA	GGAGCTCGCA	GAGTCACAAG	CGGCAGCTGC	AGCTCGAGGA	GCACGGCAGC	3960
AGTGCCAAAC	AACGCGGAGG	ACACCACCGT	CGACGAGCCC	CGGTGGTGCA	GCCGTGCATG	4020

GAGAGCGAGA	ACGAGAACAT	GCTGGCGGAG	TACGAGCAGC	GCCAGTACAC	CAGCGATTGC	4080
TGCAATAGCT	CCCGCGAGGG	CGACACCTGC	TCCTGCAGCG	AGGGATCCTG	TCTTTACGCC	4140
GAGGCGGGCG	AGCCGGCGCC	TCGTCAAATG	ACTGCTAAGA	ACACCTAA		4188

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1395 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	His	Pro	Met	His	Pro	Glu	Asn	His	Ala	Ile	Ala	Arg	Ser	Thr	Ser	1	5	10	15
Thr	Thr	Asn	Asn	Pro	Ser	Arg	Ser	Arg	Ser	Ser	Arg	Met	Trp	Leu	Leu	20	25	30	
Pro	Ala	Trp	Leu	Leu	Leu	Val	Leu	Val	Ala	Ser	Asn	Gly	Leu	Pro	Ala	35	40	45	
Val	Arg	Gly	Gln	Tyr	Gln	Ser	Pro	Arg	Ile	Ile	Glu	His	Pro	Thr	Asp	50	55	60	
Leu	Val	Val	Lys	Lys	Asn	Glu	Pro	Ala	Thr	Leu	Asn	Cys	Lys	Val	Glu	65	70	75	80
Gly	Lys	Pro	Glu	Pro	Thr	Ile	Glu	Trp	Phe	Lys	Asp	Gly	Glu	Pro	Val	85	90	95	
Ser	Thr	Asn	Glu	Lys	Lys	Ser	His	Arg	Val	Gln	Phe	Lys	Asp	Gly	Ala	100	105	110	
Leu	Phe	Phe	Tyr	Arg	Thr	Met	Gln	Gly	Lys	Lys	Glu	Gln	Asp	Gly	Gly	115	120	125	
Glu	Tyr	Trp	Cys	Val	Ala	Lys	Asn	Arg	Val	Gly	Gln	Ala	Val	Ser	Arg	130	135	140	
His	Ala	Ser	Leu	Gln	Ile	Ala	Val	Leu	Arg	Asp	Asp	Phe	Arg	Val	Glu	145	150	155	160
Pro	Lys	Asp	Thr	Arg	Val	Ala	Lys	Gly	Glu	Thr	Ala	Leu	Leu	Glu	Cys	165	170	175	
Gly	Pro	Pro	Lys	Gly	Ile	Pro	Glu	Pro	Thr	Leu	Ile	Trp	Ile	Lys	Asp	180	185	190	
Gly	Val	Pro	Leu	Asp	Asp	Leu	Lys	Ala	Met	Ser	Phe	Gly	Ala	Ser	Ser	195	200	205	
Arg	Val	Arg	Ile	Val	Asp	Gly	Gly	Asn	Leu	Leu	Ile	Ser	Asn	Val	Glu	210	215	220	
Pro	Ile	Asp	Glu	Gly	Asn	Tyr	Lys	Cys	Ile	Ala	Gln	Asn	Leu	Val	Gly	225	230	235	240
Thr	Arg	Glu	Ser	Ser	Tyr	Ala	Lys	Leu	Ile	Val	Gln	Val	Lys	Pro	Tyr	245	250	255	
Phe	Met	Lys	Glu	Pro	Lys	Asp	Gln	Val	Met	Leu	Tyr	Gly	Gln	Thr	Ala	260	265	270	
Thr	Phe	His	Cys	Ser	Val	Gly	Gly	Asp	Pro	Pro	Pro	Lys	Val	Leu	Trp	275	280	285	
Lys	Lys	Glu	Glu	Gly	Asn	Ile	Pro	Val	Ser	Arg	Ala	Arg	Ile	Leu	His	290	295	300	
Asp	Glu	Lys	Ser	Leu	Glu	Ile	Ser	Asn	Ile	Thr	Pro	Thr	Asp	Glu	Gly	305	310	315	320
Thr	Tyr	Val	Cys	Glu	Ala	His	Asn	Asn	Val	Gly	Gln	Ile	Ser	Ala	Arg	325	330	335	

Ala	Ser	Leu	Ile	Val	His	Ala	Pro	Pro	Asn	Phe	Thr	Lys	Arg	Pro	Ser		
			340						345				350				
Asn	Lys	Lys	Val	Gly	Leu	Asn	Gly	Val	Val	Gln	Leu	Pro	Cys	Met	Ala		
		355					360					365					
Ser	Gly	Asn	Pro	Pro	Pro	Ser	Val	Phe	Trp	Thr	Lys	Glu	Gly	Val	Ser		
		370					375				380						
Thr	Leu	Met	Phe	Pro	Asn	Ser	Ser	His	Gly	Arg	Gln	Tyr	Val	Ala	Ala		
385					390					395					400		
Asp	Gly	Thr	Leu	Gln	Ile	Thr	Asp	Val	Arg	Gln	Glu	Asp	Glu	Gly	Tyr		
				405					410					415			
Tyr	Val	Cys	Ser	Ala	Phe	Ser	Val	Val	Asp	Ser	Ser	Thr	Val	Arg	Val		
			420					425					430				
Phe	Leu	Gln	Val	Ser	Ser	Val	Asp	Glu	Arg	Pro	Pro	Pro	Ile	Ile	Gln		
		435					440					445					
Ile	Gly	Pro	Ala	Asn	Gln	Thr	Leu	Pro	Lys	Gly	Ser	Val	Ala	Thr	Leu		
	450					455					460						
Pro	Cys	Arg	Ala	Thr	Gly	Asn	Pro	Ser	Pro	Arg	Ile	Lys	Trp	Phe	His		
465					470					475					480		
Asp	Gly	His	Ala	Val	Gln	Ala	Gly	Asn	Arg	Tyr	Ser	Ile	Ile	Gln	Gly		
				485					490					495			
Ser	Ser	Leu	Arg	Val	Asp	Asp	Leu	Gln	Leu	Ser	Asp	Ser	Gly	Thr	Tyr		
			500					505					510				
Thr	Cys	Thr	Ala	Ser	Gly	Glu	Arg	Gly	Glu	Thr	Ser	Trp	Ala	Ala	Thr		
		515					520						525				
Leu	Thr	Val	Glu	Lys	Pro	Gly	Ser	Thr	Ser	Leu	His	Arg	Ala	Ala	Asp		
	530					535					540						
Pro	Ser	Thr	Tyr	Pro	Ala	Pro	Pro	Gly	Thr	Pro	Lys	Val	Leu	Asn	Val		
545					550					555					560		
Ser	Arg	Thr	Ser	Ile	Ser	Leu	Arg	Trp	Ala	Lys	Ser	Gln	Glu	Lys	Pro		
				565					570					575			
Gly	Ala	Val	Gly	Pro	Ile	Ile	Gly	Tyr	Thr	Val	Glu	Tyr	Phe	Ser	Pro		
			580					585					590				
Asp	Leu	Gln	Thr	Gly	Trp	Ile	Val	Ala	Ala	His	Arg	Val	Gly	Asp	Thr		
		595					600					605					
Gln	Val	Thr	Ile	Ser	Gly	Leu	Thr	Pro	Gly	Thr	Ser	Tyr	Val	Phe	Leu		
	610					615						620					
Val	Arg	Ala	Glu	Asn	Thr	Gln	Gly	Ile	Ser	Val	Pro	Ser	Gly	Leu	Ser		
625					630					635					640		
Asn	Val	Ile	Lys	Thr	Ile	Glu	Ala	Asp	Phe	Asp	Ala	Ala	Ser	Ala	Asn		
				645					650					655			
Asp	Leu	Ser	Ala	Ala	Arg	Thr	Leu	Leu	Thr	Gly	Lys	Ser	Val	Glu	Leu		
			660					665					670				
Ile	Asp	Ala	Ser	Ala	Ile	Asn	Ala	Ser	Ala	Val	Arg	Leu	Glu	Trp	Met		
		675					680						685				
Leu	His	Val	Ser	Ala	Asp	Glu	Lys	Tyr	Val	Glu	Gly	Leu	Arg	Ile	His		
	690					695					700						
Tyr	Lys	Asp	Ala	Ser	Val	Pro	Ser	Ala	Gln	Tyr	His	Ser	Ile	Thr	Val		
705					710					715					720		
Met	Asp	Ala	Ser	Ala	Glu	Ser	Phe	Val	Val	Gly	Asn	Leu	Lys	Lys	Tyr		
				725					730					735			
Thr	Lys	Tyr	Glu	Phe	Phe	Leu	Thr	Pro	Phe	Phe	Glu	Thr	Ile	Glu	Gly		
			740					745					750				
Gln	Pro	Ser	Asn	Ser	Lys	Thr	Ala	Leu	Thr	Tyr	Glu	Asp	Val	Pro	Ser		
			755					760					765				

Ala	Pro	Pro	Asp	Asn	Ile	Gln	Ile	Gly	Met	Tyr	Asn	Gln	Thr	Ala	Gly	770	775	780
Trp	Val	Arg	Trp	Thr	Pro	Pro	Pro	Ser	Gln	His	His	Asn	Gly	Asn	Leu	785	790	795
Tyr	Gly	Tyr	Lys	Ile	Glu	Val	Ser	Ala	Gly	Asn	Thr	Met	Lys	Val	Leu	805	810	815
Ala	Asn	Met	Thr	Leu	Asn	Ala	Thr	Thr	Thr	Ser	Val	Leu	Leu	Asn	Asn	820	825	830
Leu	Thr	Thr	Gly	Ala	Val	Tyr	Ser	Val	Arg	Leu	Asn	Ser	Phe	Thr	Lys	835	840	845
Ala	Gly	Asp	Gly	Pro	Tyr	Ser	Lys	Pro	Ile	Ser	Leu	Phe	Met	Asp	Pro	850	855	860
Thr	His	His	Val	His	Pro	Pro	Arg	Ala	His	Pro	Ser	Gly	Thr	His	Asp	865	870	875
Gly	Arg	His	Glu	Gly	Gln	Asp	Leu	Thr	Tyr	His	Asn	Asn	Gly	Asn	Ile	885	890	895
Pro	Pro	Gly	Asp	Ile	Asn	Pro	Thr	Thr	His	Lys	Lys	Thr	Thr	Asp	Tyr	900	905	910
Leu	Ser	Gly	Pro	Trp	Leu	Met	Val	Leu	Val	Cys	Ile	Val	Leu	Leu	Val	915	920	925
Leu	Val	Ile	Ser	Ala	Ala	Ile	Ser	Met	Val	Tyr	Phe	Lys	Arg	Lys	His	930	935	940
Gln	Met	Thr	Lys	Glu	Leu	Gly	His	Leu	Ser	Val	Val	Ser	Asp	Asn	Glu	945	950	955
Ile	Thr	Ala	Leu	Asn	Ile	Asn	Ser	Lys	Glu	Ser	Leu	Trp	Ile	Asp	His	965	970	975
His	Arg	Gly	Trp	Arg	Thr	Ala	Asp	Thr	Asp	Lys	Asp	Ser	Gly	Leu	Ser	980	985	990
Glu	Ser	Lys	Leu	Leu	Ser	His	Val	Asn	Ser	Ser	Gln	Ser	Asn	Tyr	Asn	995	1000	1005
Asn	Ser	Asp	Gly	Gly	Thr	Asp	Tyr	Ala	Glu	Val	Asp	Thr	Arg	Asn	Leu	1010	1015	1020
Thr	Thr	Phe	Tyr	Asn	Cys	Arg	Lys	Ser	Pro	Asp	Asn	Pro	Thr	Pro	Tyr	1025	1030	1035
Ala	Thr	Thr	Met	Ile	Ile	Gly	Thr	Ser	Ser	Ser	Glu	Thr	Cys	Thr	Lys	1045	1050	1055
Thr	Thr	Ser	Ile	Ser	Ala	Asp	Lys	Asp	Ser	Gly	Thr	His	Ser	Pro	Tyr	1060	1065	1070
Ser	Asp	Ala	Phe	Ala	Gly	Gln	Val	Pro	Ala	Val	Pro	Val	Val	Lys	Ser	1075	1080	1085
Asn	Tyr	Leu	Gln	Tyr	Pro	Val	Glu	Pro	Ile	Asn	Trp	Ser	Glu	Phe	Leu	1090	1095	1100
Pro	Pro	Pro	Pro	Glu	His	Pro	Pro	Pro	Ser	Ser	Thr	Tyr	Gly	Tyr	Ala	1105	1110	1115
Gln	Gly	Ser	Pro	Glu	Ser	Ser	Arg	Lys	Ser	Ser	Lys	Ser	Ala	Gly	Ser	1125	1130	1135
Gly	Ile	Ser	Thr	Asn	Gln	Ser	Ile	Leu	Asn	Ala	Ser	Ile	His	Ser	Ser	1140	1145	1150
Ser	Ser	Gly	Gly	Phe	Ser	Ala	Trp	Gly	Val	Ser	Pro	Gln	Tyr	Ala	Val	1155	1160	1165
Ala	Cys	Pro	Pro	Glu	Asn	Val	Tyr	Ser	Asn	Pro	Leu	Ser	Ala	Val	Ala	1170	1175	1180
Gly	Gly	Thr	Gln	Asn	Arg	Tyr	Gln	Ile	Thr	Pro	Thr	Asn	Gln	His	Pro	1185	1190	1195
																		1200

Pro Gln Leu Pro Ala Tyr Phe Ala Thr Thr Gly Pro Gly Gly Ala Val
1205 1210 1215
Pro Pro Asn His Leu Pro Phe Ala Thr Gln Arg His Ala Ala Ser Glu
1220 1225 1230
Tyr Gln Ala Gly Leu Asn Ala Ala Arg Cys Ala Gln Ser Arg Ala Cys
1235 1240 1245
Asn Ser Cys Asp Ala Leu Ala Thr Pro Ser Pro Met Gln Pro Pro Pro
1250 1255 1260
Pro Val Pro Val Pro Glu Gly Trp Tyr Gln Pro Val His Pro Asn Ser
1265 1270 1275 1280
His Pro Met His Pro Thr Ser Ser Asn His Gln Ile Tyr Gln Cys Ser
1285 1290 1295
Ser Glu Cys Ser Asp His Ser Arg Ser Ser Gln Ser His Lys Arg Gln
1300 1305 1310
Leu Gln Leu Glu Glu His Gly Ser Ser Ala Lys Gln Arg Gly Gly His
1315 1320 1325
His Arg Arg Arg Ala Pro Val Val Gln Pro Cys Met Glu Ser Glu Asn
1330 1335 1340
Glu Asn Met Leu Ala Glu Tyr Glu Gln Arg Gln Tyr Thr Ser Asp Cys
1345 1350 1355 1360
Cys Asn Ser Ser Arg Glu Gly Asp Thr Cys Ser Cys Ser Glu Gly Ser
1365 1370 1375
Cys Leu Tyr Ala Glu Ala Gly Glu Pro Ala Pro Arg Gln Met Thr Ala
1380 1385 1390
Lys Asn Thr
1395

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4146 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGTGAAATC	CACGCATCAT	CGAGCATCCC	ATGGACACGA	CGGTGCCAAA	AAATGATCCA	60
TTTACGTTTA	ATTGCCAGGC	CGAGGGCAAT	CCAACACCAA	CCATTCAATG	GTTTAAGGAC	120
GGTCGCGAAC	TGAAGACGGA	TACGGGTTTCG	CATCGCATAA	TGCTGCCCGC	CGGGGGTCTA	180
TTCTTTCTCA	AGGTTATCCA	CTCACGTAGA	GAGAGCGATG	CGGGCACTTA	CTGGTGCGAG	240
GCCAAAAACG	AGTTTGGAGT	GGCACGGTCC	AGGAATGCAA	CGTTGCAAGT	GGCAGTTCTC	300
CGCGACGAAT	TCCGTTTGGA	GCCGGCAAAT	ACCCGCGTGG	CCCAAGGCGA	GGTGGCCCTG	360
ATGGAATGCG	GTGCCCCCGG	AGGATCTCCG	GAGCCGCAAA	TCTCGTGGCG	CAAGAACGGC	420
CAGACCCCTGA	ATCTTGTCGG	GAACAAGCGG	ATTTCGATTG	TCGACGGTGG	CAATCTGGCC	480
ATCCAGGAAG	CCCGCCAATC	GGACGACGGA	CGCTACCAGT	GTGTGGTCAA	GAATGTGGTT	540
GGCACCCGGG	AGTCGGCCAC	CGCTTTTCTT	AAAGTGCATG	TACGTCCATT	CCTCATCCGA	600
GGACCCAGAG	ATCAGACGGC	GGTGGTGGGC	AGCTCGGTGG	TCTTCCAGTG	CCGCATCGGA	660
GGCGATCCCC	TGCCTGATGT	CCTGTGGCGA	CGCACTGCCT	CCGGCGGCAA	TATGCCACTG	720
CGTAAGTTTT	CTTGGCTTCA	TTCAGCTTCA	GGTCGTGTGC	ACGTACTTGA	GGACCGCAGT	780
CTGAAGCTGG	ACGACGTTAC	TCTGGAGGAC	ATGGGCGAGT	ACACTTGCGA	GGCGGACAAT	840
GCGGTGGGCG	GCATCACGGC	CACTGGCATC	CTCACCGTTC	ACGCTCCCCC	CAAATTTGTG	900
ATACGCCCCA	AGAATCAGCT	GGTGGAGATC	GGTGATGAAG	TGCTGTTCGA	GTGCCAAGCG	960
AATGGACATC	CCCGACCAAC	GCTCTACTGG	TCGGTGGAGG	GCAACAGCTC	CCTGCTGCTC	1020
CCCGGCTATC	GGGATGGCCG	CATGGAAGTG	ACCCTGACGC	CCGAGGGGCG	CTCGGTGCTC	1080
TCGATAGCTC	GATTTGCCCG	TGAGGATTCC	GGAAAGGTGG	TCACTTGCAA	CGCCCTGAAC	1140

GCCGTGGGCA	GCGTCAGCAG	TCGGACTGTG	GTCAGTGTGG	ATACGCAATT	CGAGCTGCCA	1200
CCGCCGATTA	TCGAACAGGG	GCCCGTGAAT	CAAACGTTGC	CCGTTAAATC	AATTGTGGTT	1260
CTGCCATGCC	GAACCTCTGGG	CACTCCAGTG	CCACAGGTCT	CTTGGTACCT	GGATGGCATA	1320
CCCATCGATG	TGCAGGAGCA	CGAGCGGCGG	AATCTTTTCGG	ACGCTGGAGC	CTTAACCATT	1380
TCGGATCTTC	AGCGCCACGA	GGATGAAGGC	TTGTACACCT	GCGTGGCCAG	CAATCGCAAC	1440
GGAAAATCCT	CTTGAGGTGG	TTACCTTCGT	CTGGACACCC	CGACAAATCC	GAATATCAAG	1500
TTCTTCAGAG	CCCCAGAACT	TTCCACCTAC	CCAGGGCCGC	CAGGAAAACC	GCAAATGGTG	1560
GAGAAGGGCG	AAAATTCGGT	GACTCTCAGC	TGGACGAGGA	GCAACAAGGT	GGGCGGCTCC	1620
AGTCTGGTGG	GCTATGTAAT	CGAGATGTTT	GGCAAAAACG	AAACGGATGG	CTGGGTGGCT	1680
GTGGGCACTA	GGGTGCAAAA	TACCACGTTT	ACCCAAACGG	GTCTGCTGCC	GGGTGTGAAT	1740
TACTTCTTTC	TAATTCGAGC	CGAGAACTCC	CATGGCTTAT	CACTGCCCAG	TCCGATGTCG	1800
GAACCCATTA	CGGTGGGAAC	GCGCTACTTC	AATAGTGGTC	TGGATCTGAG	CGAGGCTCGT	1860
GCCAGTCTGC	TGTCCGGAGA	TGTTGTGGAG	CTGAGCAACG	CCAGTGTGGT	GGACTCCACT	1920
AGCATGAAAC	TCACCTGGCA	GATCATCAAT	GGCAAAATACG	TCGAGGGCTT	CTATGTCTAT	1980
GCGAGACAGT	TGCCAAATCC	AATAGTCAAC	AATCCGGCGC	CCGTTACTAG	CAATACCAAT	2040
CCGCTGCTGG	GCTCTACATC	CACATCCGCA	TCCGCATCCG	CCTCGGCATC	GGCATTGATT	2100
TCGACAAAGC	CAAATATTGC	AGCTGCCGGC	AAACGTGATG	GGGAGACAAA	CCAGAGTGGA	2160
GGAGGAGCTC	CGACCCCACT	GAACACCAAG	TATCGCATGC	TAACGATTCT	CAATGGCGGT	2220
GGCGCCTCAT	CCTGCACCAT	CACCGGGCTC	GTCCAGTACA	CGCTGTATGA	ATTTTTTCATC	2280
GTGCCATTTT	ACAAATCCGT	CGAGGGCAAG	CCGTGCAATT	CGCGCATCGC	TCGCACCCTT	2340
GAAGATGTTT	CCTCTGAGGC	ACCATATGGA	ATGGAGGCTC	TGCTGTTGAA	CTCCTCCGCG	2400
GTCTTCCTCA	AATGGAAGGC	ACCAGAACTC	AAGGATCGGC	ATGGTGTTC	CTTGAACTAT	2460
CATGTTATAG	TCCGAGGTAT	TGACACTGCC	CACAATTTCT	CACGCATTTT	GACAAATGTC	2520
ACCATCGATG	CCGCTTCGCC	TACTCTGGTT	TTGGCCAATC	TCACCGAAGG	CGTCATGTAC	2580
ACCGTGGGCG	TGGCGGCCGG	AAATAACGCT	GGAGTTGGTC	CTTATTGTGT	CCCAGCTACT	2640
TTGCGTTTGG	ATCCCATCAC	AAAGCGACTC	GATCCGTTCA	TCAATCAGCG	GGACCATGTT	2700
AACGATGTGC	TGACGCAGCC	CTGGTTCATA	ATACTCCTGG	GCGCCATCCT	GGCCGTTCTT	2760
ATGCTGTCTT	TTGGCGCAAT	GGTCTTTGTG	AAGCGCAAGC	ACATGATGAT	GAAGCAGTCG	2820
GCCCTAAATA	CAATGCGTGG	CAATCACACG	AGCGACGTGC	TCAAAATGCC	GAGTCTATCG	2880
GCGCGCAATG	GAAACGGCTA	CTGGCTGGAC	TCCTCCACCG	GCGGAATGGT	GTGGCGTCCC	2940
TCGCCCCGGC	GCGACTCGCT	GGAGATGCAA	AAGGATCACA	TCGCCGACTA	TGCGCCGGTC	3000
TGCGGTGCCC	CCGTTCTTCC	GGCCGGCGGT	GGCACCTCTT	CCGGTGGATC	CGGTGGCGCG	3060
GGCAGCGGTG	CCAGCGGCGG	CGATGACATT	CATGGAGGAC	ACGGCAGCGA	ACGCAATCAG	3120
CAGCGGTACG	TGGGCGAGTA	CTCCAACATA	CCGACCGACT	ATGCAGAGGT	GTCCAGTTTT	3180
GGCAAGGCAC	CCAGCGAGTA	TGGTCGGCAT	GGCAACGCCT	CCCCGGCCCC	TTATGCCACC	3240
TCTTCGATCC	TGAGTCCCCA	CCAGCAGCAA	CAGCAGCAGC	AGCCGCGTTA	TCAACAGCGA	3300
CCAGTGCCCG	GCTATGGGCT	CCAGCGCCCA	ATGCACCCAC	ACTACCAGCA	GCAGCAGCAT	3360
CAGCAGCAAC	AGGCGCAGCA	GACGCACCAG	CAACACCAGG	CTCTCCAGCA	GCACCAGCAA	3420
CTGCCACCCA	GCAACATCTA	CCAGCAGATG	TCCACCACCA	GCGAGATATA	CCCCACGAAC	3480
ACGGGTCTTT	CGCGCTCTGT	CTACTCTGAG	CAGTATTACT	ACCCCAAGGA	CAAGCAGAGA	3540
CACATCCACA	TCACCGAGAA	CAAGCTGAGC	AACTGCCACA	CCTATGAGGC	GGCTCCTGGC	3600
GCCAAGCAGT	CCTCGCCGAT	ATCCTCGCAG	TTCCGCAGCG	TGAGGCGGCA	GCAGCTGCCG	3660
CCCAACTGCA	GCATCGGCAG	GGAAAGTGCC	CGCTTCAAGG	TGCTAAACAC	GGATCAGGGC	3720
AAGAACCAGC	AGAATCTCCT	GGATCTCGAC	GGCTCCTCGA	TGTGCTACAA	CGGTCTGGCA	3780
GACTCGGGCT	GCGGTGGATC	TCCCTCCCCG	ATGGCCATGC	TGATGTCGCA	CGAGGACGAG	3840
CACGCGCTGT	ACCACACGGC	GGATGGGGAT	CTGGACGACA	TGGAACGACT	GTACGTCAAG	3900
GTGGACGAGC	AGCAGCCTCC	ACAGCAGCAG	CAGCAGCTGA	TTCCCCTGGT	CCCACAGCAT	3960
CCGGCGGAAG	GTCACCTGCA	GTCCTGGCGG	AATCAGAGCA	CGCGGAGCAG	TCGGAAGAAC	4020
GGCCAGGAAT	GCATCAAGGA	ACCCAGCGAG	TTGATCTACG	CTCCGGGAAG	CGTGGCCAGC	4080
GAACGGAGCC	TCCTCAGCAA	CTCGGGTAGC	GGCACCAGCA	GCCAGCCAGC	TGGCCACAAT	4140
GTCTGA						4146

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1381 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Gly	Glu	Asn	Pro	Arg	Ile	Ile	Glu	His	Pro	Met	Asp	Thr	Thr	Val	Pro
1				5					10					15	
Lys	Asn	Asp	Pro	Phe	Thr	Phe	Asn	Cys	Gln	Ala	Glu	Gly	Asn	Pro	Thr
			20					25					30		
Pro	Thr	Ile	Gln	Trp	Phe	Lys	Asp	Gly	Arg	Glu	Leu	Lys	Thr	Asp	Thr
		35					40					45			
Gly	Ser	His	Arg	Ile	Met	Leu	Pro	Ala	Gly	Gly	Leu	Phe	Phe	Leu	Lys
	50					55					60				
Val	Ile	His	Ser	Arg	Arg	Glu	Ser	Asp	Ala	Gly	Thr	Tyr	Trp	Cys	Glu
65					70				75					80	
Ala	Lys	Asn	Glu	Phe	Gly	Val	Ala	Arg	Ser	Arg	Asn	Ala	Thr	Leu	Gln
				85				90						95	
Val	Ala	Val	Leu	Arg	Asp	Glu	Phe	Arg	Leu	Glu	Pro	Ala	Asn	Thr	Arg
			100					105					110		
Val	Ala	Gln	Gly	Glu	Val	Ala	Leu	Met	Glu	Cys	Gly	Ala	Pro	Arg	Gly
		115					120					125			
Ser	Pro	Glu	Pro	Gln	Ile	Ser	Trp	Arg	Lys	Asn	Gly	Gln	Thr	Leu	Asn
	130					135					140				
Leu	Val	Gly	Asn	Lys	Arg	Ile	Arg	Ile	Val	Asp	Gly	Gly	Asn	Leu	Ala
145					150					155				160	
Ile	Gln	Glu	Ala	Arg	Gln	Ser	Asp	Asp	Gly	Arg	Tyr	Gln	Cys	Val	Val
				165					170					175	
Lys	Asn	Val	Val	Gly	Thr	Arg	Glu	Ser	Ala	Thr	Ala	Phe	Leu	Lys	Val
			180					185					190		
His	Val	Arg	Pro	Phe	Leu	Ile	Arg	Gly	Pro	Gln	Asn	Gln	Thr	Ala	Val
	195						200					205			
Val	Gly	Ser	Ser	Val	Val	Phe	Gln	Cys	Arg	Ile	Gly	Gly	Asp	Pro	Leu
	210					215					220				
Pro	Asp	Val	Leu	Trp	Arg	Arg	Thr	Ala	Ser	Gly	Gly	Asn	Met	Pro	Leu
225					230					235				240	
Arg	Lys	Phe	Ser	Trp	Leu	His	Ser	Ala	Ser	Gly	Arg	Val	His	Val	Leu
			245						250					255	
Glu	Asp	Arg	Ser	Leu	Lys	Leu	Asp	Asp	Val	Thr	Leu	Glu	Asp	Met	Gly
		260						265					270		
Glu	Tyr	Thr	Cys	Glu	Ala	Asp	Asn	Ala	Val	Gly	Gly	Ile	Thr	Ala	Thr
		275					280					285			
Gly	Ile	Leu	Thr	Val	His	Ala	Pro	Pro	Lys	Phe	Val	Ile	Arg	Pro	Lys
	290					295						300			
Asn	Gln	Leu	Val	Glu	Ile	Gly	Asp	Glu	Val	Leu	Phe	Glu	Cys	Gln	Ala
305					310					315				320	
Asn	Gly	His	Pro	Arg	Pro	Thr	Leu	Tyr	Trp	Ser	Val	Glu	Gly	Asn	Ser
			325						330					335	
Ser	Leu	Leu	Leu	Pro	Gly	Tyr	Arg	Asp	Gly	Arg	Met	Glu	Val	Thr	Leu
			340					345					350		
Thr	Pro	Glu	Gly	Arg	Ser	Val	Leu	Ser	Ile	Ala	Arg	Phe	Ala	Arg	Glu
		355					360					365			
Asp	Ser	Gly	Lys	Val	Val	Thr	Cys	Asn	Ala	Leu	Asn	Ala	Val	Gly	Ser
	370					375						380			

Val	Ser	Ser	Arg	Thr	Val	Val	Ser	Val	Asp	Thr	Gln	Phe	Glu	Leu	Pro	385	390	395	400
Pro	Pro	Ile	Ile	Glu	Gln	Gly	Pro	Val	Asn	Gln	Thr	Leu	Pro	Val	Lys	405	410	415	
Ser	Ile	Val	Val	Leu	Pro	Cys	Arg	Thr	Leu	Gly	Thr	Pro	Val	Pro	Gln	420	425	430	
Val	Ser	Trp	Tyr	Leu	Asp	Gly	Ile	Pro	Ile	Asp	Val	Gln	Glu	His	Glu	435	440	445	
Arg	Arg	Asn	Leu	Ser	Asp	Ala	Gly	Ala	Leu	Thr	Ile	Ser	Asp	Leu	Gln	450	455	460	
Arg	His	Glu	Asp	Glu	Gly	Leu	Tyr	Thr	Cys	Val	Ala	Ser	Asn	Arg	Asn	465	470	475	480
Gly	Lys	Ser	Ser	Trp	Ser	Gly	Tyr	Leu	Arg	Leu	Asp	Thr	Pro	Thr	Asn	485	490	495	
Pro	Asn	Ile	Lys	Phe	Phe	Arg	Ala	Pro	Glu	Leu	Ser	Thr	Tyr	Pro	Gly	500	505	510	
Pro	Pro	Gly	Lys	Pro	Gln	Met	Val	Glu	Lys	Gly	Glu	Asn	Ser	Val	Thr	515	520	525	
Leu	Ser	Trp	Thr	Arg	Ser	Asn	Lys	Val	Gly	Gly	Ser	Ser	Leu	Val	Gly	530	535	540	
Tyr	Val	Ile	Glu	Met	Phe	Gly	Lys	Asn	Glu	Thr	Asp	Gly	Trp	Val	Ala	545	550	555	560
Val	Gly	Thr	Arg	Val	Gln	Asn	Thr	Thr	Phe	Thr	Gln	Thr	Gly	Leu	Leu	565	570	575	
Pro	Gly	Val	Asn	Tyr	Phe	Phe	Leu	Ile	Arg	Ala	Glu	Asn	Ser	His	Gly	580	585	590	
Leu	Ser	Leu	Pro	Ser	Pro	Met	Ser	Glu	Pro	Ile	Thr	Val	Gly	Thr	Arg	595	600	605	
Tyr	Phe	Asn	Ser	Gly	Leu	Asp	Leu	Ser	Glu	Ala	Arg	Ala	Ser	Leu	Leu	610	615	620	
Ser	Gly	Asp	Val	Val	Glu	Leu	Ser	Asn	Ala	Ser	Val	Val	Asp	Ser	Thr	625	630	635	640
Ser	Met	Lys	Leu	Thr	Trp	Gln	Ile	Ile	Asn	Gly	Lys	Tyr	Val	Glu	Gly	645	650	655	
Phe	Tyr	Val	Tyr	Ala	Arg	Gln	Leu	Pro	Asn	Pro	Ile	Val	Asn	Asn	Pro	660	665	670	
Ala	Pro	Val	Thr	Ser	Asn	Thr	Asn	Pro	Leu	Leu	Gly	Ser	Thr	Ser	Thr	675	680	685	
Ser	Ala	Ser	Ala	Ser	Ala	Ser	Ala	Ser	Ala	Leu	Ile	Ser	Thr	Lys	Pro	690	695	700	
Asn	Ile	Ala	Ala	Ala	Gly	Lys	Arg	Asp	Gly	Glu	Thr	Asn	Gln	Ser	Gly	705	710	715	720
Gly	Gly	Ala	Pro	Thr	Pro	Leu	Asn	Thr	Lys	Tyr	Arg	Met	Leu	Thr	Ile	725	730	735	
Leu	Asn	Gly	Gly	Gly	Ala	Ser	Ser	Cys	Thr	Ile	Thr	Gly	Leu	Val	Gln	740	745	750	
Tyr	Thr	Leu	Tyr	Glu	Phe	Phe	Ile	Val	Pro	Phe	Tyr	Lys	Ser	Val	Glu	755	760	765	
Gly	Lys	Pro	Ser	Asn	Ser	Arg	Ile	Ala	Arg	Thr	Leu	Glu	Asp	Val	Pro	770	775	780	
Ser	Glu	Ala	Pro	Tyr	Gly	Met	Glu	Ala	Leu	Leu	Leu	Asn	Ser	Ser	Ala	785	790	795	800
Val	Phe	Leu	Lys	Trp	Lys	Ala	Pro	Glu	Leu	Lys	Asp	Arg	His	Gly	Val	805	810	815	

Leu	Leu	Asn	Tyr	His	Val	Ile	Val	Arg	Gly	Ile	Asp	Thr	Ala	His	Asn		
			820						825					830			
Phe	Ser	Arg	Ile	Leu	Thr	Asn	Val	Thr	Ile	Asp	Ala	Ala	Ser	Pro	Thr		
		835						840				845					
Leu	Val	Leu	Ala	Asn	Leu	Thr	Glu	Gly	Val	Met	Tyr	Thr	Val	Gly	Val		
	850					855					860						
Ala	Ala	Gly	Asn	Asn	Ala	Gly	Val	Gly	Pro	Tyr	Cys	Val	Pro	Ala	Thr		
865					870					875					880		
Leu	Arg	Leu	Asp	Pro	Ile	Thr	Lys	Arg	Leu	Asp	Pro	Phe	Ile	Asn	Gln		
			885						890					895			
Arg	Asp	His	Val	Asn	Asp	Val	Leu	Thr	Gln	Pro	Trp	Phe	Ile	Ile	Leu		
			900						905					910			
Leu	Gly	Ala	Ile	Leu	Ala	Val	Leu	Met	Leu	Ser	Phe	Gly	Ala	Met	Val		
	915						920					925					
Phe	Val	Lys	Arg	Lys	His	Met	Met	Met	Lys	Gln	Ser	Ala	Leu	Asn	Thr		
	930					935					940						
Met	Arg	Gly	Asn	His	Thr	Ser	Asp	Val	Leu	Lys	Met	Pro	Ser	Leu	Ser		
945					950					955					960		
Ala	Arg	Asn	Gly	Asn	Gly	Tyr	Trp	Leu	Asp	Ser	Ser	Thr	Gly	Gly	Met		
			965						970					975			
Val	Trp	Arg	Pro	Ser	Pro	Gly	Gly	Asp	Ser	Leu	Glu	Met	Gln	Lys	Asp		
			980					985					990				
His	Ile	Ala	Asp	Tyr	Ala	Pro	Val	Cys	Gly	Ala	Pro	Gly	Ser	Pro	Ala		
	995					1000						1005					
Gly	Gly	Gly	Thr	Ser	Ser	Gly	Gly	Ser	Gly	Gly	Ala	Gly	Ser	Gly	Ala		
	1010					1015					1020						
Ser	Gly	Gly	Asp	Asp	Ile	His	Gly	Gly	His	Gly	Ser	Glu	Arg	Asn	Gln		
1025					1030					1035					1040		
Gln	Arg	Tyr	Val	Gly	Glu	Tyr	Ser	Asn	Ile	Pro	Thr	Asp	Tyr	Ala	Glu		
			1045						1050					1055			
Val	Ser	Ser	Phe	Gly	Lys	Ala	Pro	Ser	Glu	Tyr	Gly	Arg	His	Gly	Asn		
			1060					1065					1070				
Ala	Ser	Pro	Ala	Pro	Tyr	Ala	Thr	Ser	Ser	Ile	Leu	Ser	Pro	His	Gln		
	1075						1080					1085					
Gln	Gln	Gln	Gln	Gln	Gln	Pro	Arg	Tyr	Gln	Gln	Arg	Pro	Val	Pro	Gly		
	1090					1095					1100						
Tyr	Gly	Leu	Gln	Arg	Pro	Met	His	Pro	His	Tyr	Gln	Gln	Gln	Gln	His		
1105					1110					1115					1120		
Gln	Gln	Gln	Gln	Ala	Gln	Gln	Thr	His	Gln	Gln	His	Gln	Ala	Leu	Gln		
			1125					1130					1135				
Gln	His	Gln	Gln	Leu	Pro	Pro	Ser	Asn	Ile	Tyr	Gln	Gln	Met	Ser	Thr		
			1140					1145					1150				
Thr	Ser	Glu	Ile	Tyr	Pro	Thr	Asn	Thr	Gly	Pro	Ser	Arg	Ser	Val	Tyr		
	1155						1160					1165					
Ser	Glu	Gln	Tyr	Tyr	Tyr	Pro	Lys	Asp	Lys	Gln	Arg	His	Ile	His	Ile		
	1170					1175				1180							
Thr	Glu	Asn	Lys	Leu	Ser	Asn	Cys	His	Thr	Tyr	Glu	Ala	Ala	Pro	Gly		
1185					1190					1195				1200			
Ala	Lys	Gln	Ser	Ser	Pro	Ile	Ser	Ser	Gln	Phe	Ala	Ser	Val	Arg	Arg		
			1205						1210				1215				
Gln	Gln	Leu	Pro	Pro	Asn	Cys	Ser	Ile	Gly	Arg	Glu	Ser	Ala	Arg	Phe		
			1220					1225					1230				
Lys	Val	Leu	Asn	Thr	Asp	Gln	Gly	Lys	Asn	Gln	Gln	Asn	Leu	Leu	Asp		
			1235				1240					1245					

Leu Asp Gly Ser Ser Met Cys Tyr Asn Gly Leu Ala Asp Ser Gly Cys
 1250 1255 1260
 Gly Gly Ser Pro Ser Pro Met Ala Met Leu Met Ser His Glu Asp Glu
 1265 1270 1275 1280
 His Ala Leu Tyr His Thr Ala Asp Gly Asp Leu Asp Asp Met Glu Arg
 1285 1290 1295
 Leu Tyr Val Lys Val Asp Glu Gln Gln Pro Pro Gln Gln Gln Gln Gln
 1300 1305 1310
 Leu Ile Pro Leu Val Pro Gln His Pro Ala Glu Gly His Leu Gln Ser
 1315 1320 1325
 Trp Arg Asn Gln Ser Thr Arg Ser Ser Arg Lys Asn Gly Gln Glu Cys
 1330 1335 1340
 Ile Lys Glu Pro Ser Glu Leu Ile Tyr Ala Pro Gly Ser Val Ala Ser
 1345 1350 1355 1360
 Glu Arg Ser Leu Leu Ser Asn Ser Gly Ser Gly Thr Ser Ser Gln Pro
 1365 1370 1375
 Ala Gly His Asn Val
 1380

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3894 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATGTACTATC TAGGTTTTTA CCACACTCAC ACACACACAC ACACATACAT AAATTTTGAT	60
AAAATTCCTA ATGCCTCAAA TCTCGCTCCC GTGATAATCG AACATCCCAT CGATGTGGTG	120
GTATCTAGGG GATCGCCAGC AACCTCAAC TGTGGTGCAA AGCCATCTAC CGCCAAAATC	180
ACATGGTACA AGGATGGACA GCCCGTAATC ACGAATAAGG AGCAAGTGAA CAGCCACCGG	240
ATTGTTCTCG ACACGGGATC CCTGTTTCTT CTGAAAGTGA ATAGTGGAAA AAACGGAAAA	300
GACAGCGATG CGGGAGCGTA CTATTGTGTG GCCAGCAACG AGCACGGAGA AGTGAAGTCG	360
AACGAAGGAT CGTTAAAAAT GGCGATGCTT CGCGAAGACT TTCGAGTTCG GCCAAGAACA	420
GTTTCAGGCTC TTGGTGGAGA GATGGCCGTT CTGGAATGCA GTCCGCCACG TGGATTCCCG	480
GAGCCGGTTG TGAGCTGGCG GAAAGACGAC AAAGAGCTCC GAATTCAAGA CATGCCACGA	540
TACACTCTAC ACTCTGACGG AAACCTCATC ATTGATCCGG TCGATCGAAG CGATTCTGGT	600
ACTTATCAGT GTGTTGCCAA CAACATGGTC GGAGAACGGG TGTCCAATCC CGCAAGATTG	660
AGTGTCTTTG AGAAACCAAA GTTTGAGCAA GAACCAAGG ACATGACGGT CGACGTCGGA	720
GCCGCAGTGC TGTTTGATTG TCGTGTGACT GGAGATCCTC AACCACAAAT TACGTGGAAA	780
CGCAAAAATG AGCCGATGCC AGTTACACGT GCATACATTG CCAAGGATAA TCGGGGGTTG	840
AGAATCGAAA GAGTTCAACC ATCAGACGAA GGTGAATACG TTTGCTATGC ACGAAATCCA	900
GCGGGAACCTC TTGAAGCATC TGCACATCTT CGTGTCCAGG CACCTCCATC CTTCCAGACA	960
AAACCAGCAG ACCAGTCAGT TCCAGCTGGA GGCACGGCAA CTTTTGAATG CACCTTGGTC	1020
GGTCAACCGA GTCCCGCCTA TTTTGGAGC AAGGAAGGCC AACAGGATCT TCTTTTCCCA	1080
AGTTATGTGT CCGCTGATGG TAGAACGAAA GTTTCACCAA CTGGAACATT GACAATTGAG	1140
GAAGTTCGTC AAGTTGATGA GGGAGCTTAT GTGTGCGCTG GAATGAACTC GGCAGGAAGC	1200
TCGTTGAGCA AGGCAGCTTT GAAAGCAACA TTTGAAACCA AAGGCCGTGT CCAAAAAAAA	1260
AAGAGCAAAA TGGGCAACA GAAACAAAA AATGTTCAAT CAATTATCAA ATATTTAATT	1320
TCAGCCGTGA CCGGAAACAC ACCCGCCAAA CCACCACCAA CAATCGAGCA TGGTCATCAA	1380
AATCAGACCC TTATGGTTGG ATCATCAGCC ATCCTTCCAT GTCAGGCTAG CGGAAAACCA	1440
ACTCCAGGAA TATCATGGCT CAGGGATGGG CTACCTATTG ACATTACAGA TAGTCGTATC	1500
AGTCAACATT CAACGGGAAG TCTACATATT GCCGATTTAA AGAAACCTGA CACCGGAGTT	1560
TACACTTGCA TTGCGAAGAA CGAGGATGGA GAGTCAACAT GGTCGGCATC TCTGACTGTT	1620

50					55					60					
Asp	Gly	Gln	Pro	Val	Ile	Thr	Asn	Lys	Glu	Gln	Val	Asn	Ser	His	Arg
65					70					75					80
Ile	Val	Leu	Asp	Thr	Gly	Ser	Leu	Phe	Leu	Leu	Lys	Val	Asn	Ser	Gly
				85					90					95	
Lys	Asn	Gly	Lys	Asp	Ser	Asp	Ala	Gly	Ala	Tyr	Tyr	Cys	Val	Ala	Ser
			100					105					110		
Asn	Glu	His	Gly	Glu	Val	Lys	Ser	Asn	Glu	Gly	Ser	Leu	Lys	Leu	Ala
		115					120					125			
Met	Leu	Arg	Glu	Asp	Phe	Arg	Val	Arg	Pro	Arg	Thr	Val	Gln	Ala	Leu
	130					135					140				
Gly	Gly	Glu	Met	Ala	Val	Leu	Glu	Cys	Ser	Pro	Pro	Arg	Gly	Phe	Pro
145					150					155					160
Glu	Pro	Val	Val	Ser	Trp	Arg	Lys	Asp	Asp	Lys	Glu	Leu	Arg	Ile	Gln
				165					170					175	
Asp	Met	Pro	Arg	Tyr	Thr	Leu	His	Ser	Asp	Gly	Asn	Leu	Ile	Ile	Asp
			180					185					190		
Pro	Val	Asp	Arg	Ser	Asp	Ser	Gly	Thr	Tyr	Gln	Cys	Val	Ala	Asn	Asn
		195					200					205			
Met	Val	Gly	Glu	Arg	Val	Ser	Asn	Pro	Ala	Arg	Leu	Ser	Val	Phe	Glu
	210					215					220				
Lys	Pro	Lys	Phe	Glu	Gln	Glu	Pro	Lys	Asp	Met	Thr	Val	Asp	Val	Gly
225					230					235					240
Ala	Ala	Val	Leu	Phe	Asp	Cys	Arg	Val	Thr	Gly	Asp	Pro	Gln	Pro	Gln
				245					250					255	
Ile	Thr	Trp	Lys	Arg	Lys	Asn	Glu	Pro	Met	Pro	Val	Thr	Arg	Ala	Tyr
			260					265					270		
Ile	Ala	Lys	Asp	Asn	Arg	Gly	Leu	Arg	Ile	Glu	Arg	Val	Gln	Pro	Ser
		275					280					285			
Asp	Glu	Gly	Glu	Tyr	Val	Cys	Tyr	Ala	Arg	Asn	Pro	Ala	Gly	Thr	Leu
	290					295					300				
Glu	Ala	Ser	Ala	His	Leu	Arg	Val	Gln	Ala	Pro	Pro	Ser	Phe	Gln	Thr
305					310					315					320
Lys	Pro	Ala	Asp	Gln	Ser	Val	Pro	Ala	Gly	Gly	Thr	Ala	Thr	Phe	Glu
				325					330					335	
Cys	Thr	Leu	Val	Gly	Gln	Pro	Ser	Pro	Ala	Tyr	Phe	Trp	Ser	Lys	Glu
			340					345					350		
Gly	Gln	Gln	Asp	Leu	Leu	Phe	Pro	Ser	Tyr	Val	Ser	Ala	Asp	Gly	Arg
		355					360					365			
Thr	Lys	Val	Ser	Pro	Thr	Gly	Thr	Leu	Thr	Ile	Glu	Glu	Val	Arg	Gln
	370					375					380				
Val	Asp	Glu	Gly	Ala	Tyr	Val	Cys	Ala	Gly	Met	Asn	Ser	Ala	Gly	Ser
385					390					395					400
Ser	Leu	Ser	Lys	Ala	Ala	Leu	Lys	Ala	Thr	Phe	Glu	Thr	Lys	Gly	Arg
				405					410					415	

				485					490					495	
Asp	Ser	Arg	Ile	Ser	Gln	His	Ser	Thr	Gly	Ser	Leu	His	Ile	Ala	Asp
			500					505					510		
Leu	Lys	Lys	Pro	Asp	Thr	Gly	Val	Thr	Tyr	Cys	Ile	Ala	Lys	Asn	Glu
		515					520					525			
Asp	Gly	Glu	Ser	Thr	Trp	Ser	Ala	Ser	Leu	Thr	Val	Glu	Asp	His	Thr
	530					535					540				
Ser	Asn	Ala	Gln	Phe	Val	Arg	Met	Pro	Asp	Pro	Ser	Asn	Phe	Pro	Ser
545					550					555					560
Ser	Pro	Thr	Gln	Pro	Ile	Ile	Val	Asn	Val	Thr	Asp	Thr	Glu	Val	Glu
			565						570					575	
Leu	His	Trp	Asn	Ala	Pro	Ser	Thr	Ser	Gly	Ala	Gly	Pro	Ile	Thr	Gly
			580					585					590		
Tyr	Ile	Ile	Gln	Tyr	Tyr	Ser	Pro	Asp	Leu	Gly	Gln	Thr	Trp	Phe	Asn
		595					600					605			
Ile	Pro	Asp	Tyr	Val	Ala	Ser	Thr	Glu	Tyr	Arg	Ile	Lys	Gly	Leu	Lys
	610					615					620				
Pro	Ser	His	Ser	Tyr	Met	Phe	Val	Ile	Arg	Ala	Glu	Asn	Glu	Lys	Gly
625					630					635					640
Ile	Gly	Thr	Pro	Ser	Val	Ser	Ser	Ala	Leu	Val	Thr	Thr	Ser	Lys	Pro
			645						650					655	
Ala	Ala	Gln	Val	Ala	Leu	Ser	Asp	Lys	Asn	Lys	Met	Asp	Met	Ala	Ile
		660						665					670		
Ala	Glu	Lys	Arg	Leu	Thr	Ser	Glu	Gln	Leu	Ile	Lys	Leu	Glu	Glu	Val
		675					680					685			
Lys	Thr	Ile	Asn	Ser	Thr	Ala	Val	Arg	Leu	Phe	Trp	Lys	Lys	Arg	Lys
	690					695					700				
Leu	Glu	Glu	Leu	Ile	Asp	Gly	Tyr	Tyr	Ile	Lys	Trp	Arg	Gly	Pro	Pro
705					710					715					720
Arg	Thr	Asn	Asp	Asn	Gln	Tyr	Val	Asn	Val	Thr	Ser	Pro	Ser	Thr	Glu
			725						730					735	
Asn	Tyr	Val	Val	Ser	Asn	Leu	Met	Pro	Phe	Thr	Asn	Tyr	Glu	Phe	Phe
		740						745					750		
Val	Ile	Pro	Tyr	His	Ser	Gly	Val	His	Ser	Ile	His	Gly	Ala	Pro	Ser
		755					760					765			
Asn	Ser	Met	Asp	Val	Leu	Thr	Ala	Glu	Ala	Pro	Pro	Ser	Leu	Pro	Pro
		770				775					780				
Glu	Asp	Val	Arg	Ile	Arg	Met	Leu	Asn	Leu	Thr	Thr	Leu	Arg	Ile	Ser
785					790					795					800
Trp	Lys	Ala	Pro	Lys	Ala	Asp	Gly	Ile	Asn	Gly	Ile	Leu	Lys	Gly	Phe
			805						810					815	
Gln	Ile	Val	Ile	Val	Gly	Gln	Ala	Pro	Asn	Asn	Asn	Arg	Asn	Ile	Thr
		820						825					830		
Thr	Asn	Glu	Arg	Ala	Ala	Ser	Val	Thr	Leu	Phe	His	Leu	Val</		

915	920	925
Asn Ser Asp Gly Lys Asp Arg Ser Phe Ile Lys Ile Asn Asp Gly Ser		
930	935	940
Val His Met Ala Ser Asn Asn Leu Trp Asp Val Ala Gln Asn Pro Asn		
945	950	955
Gln Asn Pro Met Tyr Asn Thr Ala Gly Arg Met Thr Met Asn Asn Arg		
965	970	975
Asn Gly Gln Ala Leu Tyr Ser Leu Thr Pro Asn Ala Gln Asp Phe Phe		
980	985	990
Asn Asn Cys Asp Asp Tyr Ser Gly Thr Met His Arg Pro Gly Ser Glu		
995	1000	1005
His His Tyr His Tyr Ala Gln Leu Thr Gly Gly Pro Gly Asn Ala Met		
1010	1015	1020
Ser Thr Phe Tyr Gly Asn Gln Tyr His Asp Asp Pro Ser Pro Tyr Ala		
1025	1030	1035
Thr Thr Thr Leu Val Leu Ser Asn Gln Gln Pro Ala Trp Leu Asn Asp		
1045	1050	1055
Lys Met Leu Arg Ala Pro Ala Met Pro Thr Asn Pro Val Pro Pro Glu		
1060	1065	1070
Pro Pro Ala Arg Tyr Ala Asp His Thr Ala Gly Arg Arg Ser Arg Ser		
1075	1080	1085
Ser Arg Ala Ser Asp Gly Arg Gly Thr Leu Asn Gly Gly Leu His His		
1090	1095	1100
Arg Thr Ser Gly Ser Gln Arg Ser Asp Ser Pro Pro His Thr Asp Val		
1105	1110	1115
Ser Tyr Val Gln Leu His Ser Ser Asp Gly Thr Gly Ser Ser Lys Glu		
1125	1130	1135
Arg Thr Gly Glu Arg Arg Thr Pro Pro Asn Lys Thr Leu Met Asp Phe		
1140	1145	1150
Ile Pro Pro Pro Pro Ser Asn Pro Pro Pro Pro Gly Gly His Val Tyr		
1155	1160	1165
Asp Thr Ala Thr Arg Arg Gln Leu Asn Arg Gly Ser Thr Pro Arg Glu		
1170	1175	1180
Asp Thr Tyr Asp Ser Val Scr Asp Gly Ala Phe Ala Arg Val Asp Val		
1185	1190	1195
Asn Ala Arg Pro Thr Ser Arg Asn Arg Asn Leu Gly Gly Arg Pro Leu		
1205	1210	1215
Lys Gly Lys Arg Asp Asp Asp Ser Gln Arg Ser Ser Leu Met Met Asp		
1220	1225	1230
Asp Asp Gly Gly Ser Ser Glu Ala Asp Gly Glu Asn Ser Glu Gly Asp		
1235	1240	1245
Val Pro Arg Gly Gly Val Arg Lys Ala Val Pro Arg Met Gly Ile Ser		
1250	1255	1260
Ala Ser Thr Leu Ala His Ser Cys Tyr Gly Thr Asn Gly Thr Ala Gln		
1265	1270	1275
Arg Phe Arg Ser Ile Pro Arg Asn Asn Gly Ile Val Thr Gln Glu Gln		
1285	1290	1295
Thr		

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4956 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ATGAAATGGA	AACATGTTCC	TTTTTTGGTC	ATGATATCAC	TCCTCAGCTT	ATCCCCAAAT	60
CACCTGTTTC	TGGCCAGCT	TATTCCAGAC	CCTGAAGATG	TAGAGAGGGG	GAACGACCAC	120
GGGACGCCAA	TCCCCACCTC	TGATAACGAT	GACAATTCGC	TGGGCTATAC	AGGCTCCCGT	180
CTTCGTCAGG	AAGATTTTCC	ACCTCGCATT	GTTGAACACC	CTTCAGACCT	GATTGTCTCA	240
AAAGGAGAAC	CTGCAACTTT	GAAGTGCAAA	GCTGAAGGCC	GCCCCACACC	CACTATTGAA	300
TGGTACAAAG	GGGGAGAGAG	AGTGGAGACA	GACAAAGATG	ACCCTCGCTC	ACACCGAATG	360
TTGCTGCCGA	GTGGATCTTT	ATTTTCTTA	CGTATAGTAC	ATGGACGGAA	AAGTAGACCT	420
GATGAAGGAG	TCTATGTCTG	TGTAGCAAGG	AATTACCTTG	GAGAGGCTGT	GAGCCACAAT	480
GCATCGCTGG	AAGTAGCCAT	ACTTCGGGAT	GAATTCAGAC	AAAACCTTC	GGATGTCATG	540
GTTGCAGTAG	GAGAGCCTGC	AGTAATGGAA	TGCCAACCTC	CACGAGGCCA	TCCTGAGCCC	600
ACCATTTTCAT	GGAAGAAAGA	TGGCTCTCCA	CTGGATGATA	AAGATGAAAG	AATAACTATA	660
CGAGGAGGAA	AGCTCATGAT	CACTTACACC	CGTAAAAGTG	ACGCTGGCAA	ATATGTTTGT	720
GTTGGTACCA	ATATGGTTGG	GGAACGTGAG	AGTGAAGTAG	CCGAGCTGAC	TGTCTTAGAG	780
AGACCATCAT	TTGTGAAGAG	ACCCAGTAAC	TTGGCAGTAA	CTGTGGATGA	CAGTGCAGAA	840
TTTAAATGTG	AGGCCCGAGG	TGACCCTGTA	CCTACAGTAC	GATGGAGGAA	AGATGATGGA	900
GAGCTGCCCA	AATCCAGATA	TGAAATCCGA	GATGATCATA	CCTTGAAAAT	TAGGAAGGTG	960
ACAGCTGGTG	ACATGGGTTT	ATACACTTGT	GTTGCAGAAA	ATATGGTGGG	CAAAGCTGAA	1020
GCATCTGCTA	CTCTGACTGT	TCAAGAACCT	CCACATTTTG	TTGTGAAACC	CCGTGACCAG	1080
GTTGTTGCTT	TGGGACGGAC	TGTAACTTTT	CAGTGTGAAG	CAACCGGAAA	TCCTCAACCA	1140
GCTATTTTCT	GGAGGAGAGA	AGGGAGTCAG	AATCTACTTT	TCTCATATCA	ACCACCACAG	1200
TCATCCAGCC	GATTTTCAGT	CTCCCAGACT	GGCGACCTCA	CAATTACTAA	TGTCCAGCGA	1260
TCTGATGTTG	GTTATTACAT	CTGCCAGACT	TTAAATGTTG	CTGGAAGCAT	CATCACAAG	1320
GCATATTTGG	AAGTTACAGA	TGTGATTGCA	GATCGGCCTC	CCCCAGTTAT	TCGACAAGGT	1380
CCTGTGAATC	AGACTGTAGC	CGTGGATGGC	ACTTTCGTCC	TCAGCTGTGT	GGCCACAGGC	1440
AGTCCAGTGC	CCACCATTCT	GTGGAGAAAAG	GATGGAGTCC	TCGTTTCAAC	CCAAGACTCT	1500
CGAATCAAAC	AGTTGGAGAA	TGGAGTACTG	CAGATCCGAT	ATGCTAAGCT	GGGTGATACT	1560
GGTCGGTACA	CCTGCATTGC	ATCAACCCCC	AGTGGTGAAG	CAACATGGAG	TGCTTACATT	1620
GAAGTTCAAG	AATTTGGAGT	TCCAGTTTCAG	CCTCCAAGAC	CTACTGACCC	AAATTTAATC	1680
CCTAGTGCCC	CATCAAAACC	TGAAGTGACA	GATGTCAGCA	GAAATACAGT	CACATTATCG	1740
TGGCAACCAA	ATTTGAATTC	AGGAGCAACT	CCAACATCTT	ATATTATAGA	AGCCTTCAGC	1800
CATGCATCTG	GTAGCAGCTG	GCAGACCGTA	GCAGAGAATG	TGAAAACAGA	AACATCTGCC	1860
ATTAAAGGAC	TCAAACCTAA	TGCAATTTAC	CTTTTCCTTG	TGAGGGCAGC	TAATGCATAT	1920
GGAATTAGTG	ATCCAAGCCA	AATATCAGAT	CCAGTGAAAA	CACAAGATGT	CCTACCAACA	1980
AGTCAGGGGG	TGGACCACAA	GCAGGTCCAG	AGAGAGCTGG	GAAATGCTGT	TCTGCACCTC	2040
CACAACCCCA	CCGTCTTTTC	TTCTCTTTCC	ATCGAAGTGC	ACTGGACAGT	AGATCAACAG	2100
TCTCAGTATA	TACAAGGATA	TAAAATTTCT	TATCGGCCAT	CTGGAGCCAA	CCACGGAGAA	2160
TCAGACTGGT	TAGTTTTTGA	AGTGAGGACG	CCAGCCAAAA	ACAGTGTGGT	AATCCCTGAT	2220
CTCAGAAAGG	GAGTCAACTA	TGAAATTAAG	GCTCGCCCTT	TTTTTAATGA	ATTTCAAGGA	2280
GCAGATAGTG	AAATCAAATT	TGCCAAAACC	CTGGAAGAAG	CACCCAGTGC	CCCACCCCAA	2340
GGTGTAACTG	TATCCAAGAA	TGATGGAAAC	GGAAGTGCAA	TTCTAGTTAG	TTGGCAGCCA	2400
CCTCCAGAAG	ACACTCAAAA	TGGAATGGTC	CAAGAGTATA	AGGTTTGGTG	TCTGGGCAAT	2460
GAAACTCGAT	ACCACATCAA	CAAAACAGTG	GATGGTTCCA	CCTTTTCCGT	GGTCATTCCC	2520
TTTCTTGTTT	CTGGAAATCCG	ATACAGTGTG	GAAGTGGCAG	CCAGCACTGG	GGCTGGGTCT	2580
GGGGTAAAGA	GTGAGCCTCA	GTTTCATCCAG	CTGGATGCCC	ATGGAAAACC	TGTGTCACTT	2640
GAGGACCAAG	TCAGCCTCGC	TCAGCAGATT	TCAGATGTGG	TGAAGCAGCC	GGCCTTCATA	2700
GCAGGTATTG	GAGCAGCTTG	TTGGATCATC	CTCATGGTCT	TCAGCATCTG	GCTTTATCGA	2760
CACCGCAAGA	AGAGAAACGG	ACTTACTAGT	ACCTACGCGG	GTATCAGAAA	AGTCCCGTCT	2820
TTTACCTTCA	CACCAACAGT	AACTTACCAG	AGAGGAGGCG	AAGCTGTCAG	CAGTGGAGGG	2880
AGGCCTGGAC	TTCTCAACAT	CAGTGAACCT	GCCGCGCAGC	CATGGCTGGC	AGACACGTGG	2940
CCTAATACTG	GCAACAACCA	CAATGACTGC	TCCATCAGCT	GCTGCACGGC	AGGCAATGGA	3000
AACAGCGACA	GCAACCTCAC	TACCTACAGT	CGCCCAGCTG	ATTGTATAGC	AAATTATAAC	3060

AACCAACTGG	ATAACAAACA	AACAAATCTG	ATGCTCCCTG	AGTCAACTGT	TTATGGTGAT	3120
GTGGACCTTA	GTAACAAAAT	CAATGAGATG	AAAACCTTCA	ATAGCCCCAA	TCTGAAGGAT	3180
GGGCGTTTTG	TCAATCCATC	AGGGCAGCCT	ACTCCTTACG	CCACCACTCA	GCTCATCCAG	3240
TCAAACCTCA	GCAACAACAT	GAACAATGGC	AGCGGGGACT	CTGGCGAGAA	GCACTGGAAA	3300
CCACTGGGAC	AGCAGAAACA	AGAAGTGGCA	CCAGTTCAGT	ACAACATCGT	GGAGCAAAAC	3360
AAGCTGAACA	AAGATTATCG	AGCAAATGAC	ACAGTTCCTC	CAACTATCCC	ATACAACCAA	3420
TCATACGACC	AGAACACAGG	AGGATCCTAC	AACAGCTCAG	ACCGGGGCAG	TAGTACATCT	3480
GGGAGTCAGG	GGCACAAGAA	AGGGGCAAGA	ACACCCAAGG	TACCAAAACA	GGGTGGCATG	3540
AACTGGGCAG	ACCTGCTTCC	TCCTCCCCCA	GCACATCCTC	CTCCACACAG	CAATAGCGAA	3600
GAGTACAACA	TTTCTGTAGA	TGAAAGCTAT	GACCAAGAAA	TGCCATGTCC	CGTGCCACCA	3660
GCAAGGATGT	ATTTGCAACA	AGATGAATTA	GAAGAGGAGG	AAGATGAACG	AGGCCCCACT	3720
CCCCCTGTTC	GGGGAGCAGC	TTCTTCTCCA	GCTGCCGTGT	CCTATAGCCA	TCAGTCCACT	3780
GCCACTCTGA	CTCCCTCCCC	ACAGGAAGAA	CTCCAGCCCA	TGTTACAGGA	TTGTCCAGAG	3840
GAGACTGGCC	ACATGCAGCA	CCAGCCCGAC	AGGAGACGGC	AGCCTGTGAG	TCCTCCTCCA	3900
CCACCACGGC	CGATCTCCCC	TCCACATACC	TATGGCTACA	TTTCAGGACC	CCTGGTCTCA	3960
GATATGGATA	CGGATGCGCC	AGAAGAGGAA	GAAGACGAAG	CCGACATGGA	GGTAGCCAAG	4020
ATGCAAACCA	GAAGGCTTTT	GTTACGTGGG	CTTGAGCAGA	CACCTGCCTC	CAGTGTGGGG	4080
GACCTGGAGA	GCTCTGTCAC	GGGGTCCATG	ATCAACGGCT	GGGGCTCAGC	CTCAGAGGAG	4140
GACAACATTT	CCAGCGGACG	CTCCAGTGTT	AGTTCTTCGG	ACGGCTCCTT	TTTCACTGAT	4200
GCTGACTTTG	CCCAGGCAGT	CGCAGCAGCG	GCAGAGTATG	CTGGTCTGAA	AGTAGCACGA	4260
CGGCAAATGC	AGGATGCTGC	TGGCCGTCGA	CATTTTCATG	CGTCTCAGTG	CCCTAGGCCC	4320
ACAAGTCCCG	TGTCTACAGA	CAGCAACATG	AGTGCCGCCG	TAATGCAGAA	AACCAGACCA	4380
GCCAAGAAAC	TGAAACACCA	GCCAGGACAT	CTGCGCAGAG	AAACCTACAC	AGATGATCTT	4440
CCACCACCTC	CTGTGCCGCC	ACCTGCTATA	AAGTCACCTA	CTGCCCAATC	CAAGACACAG	4500
CTGGAAGTAC	GACCTGTAGT	GGTGCCAAAA	CTCCCTTCTA	TGGATGCAAG	AACAGACAGA	4560
TCATCAGACA	GAAAAGGAAG	CAGTTACAAG	GGGAGAGAAG	TGTTGGATGG	AAGACAGGTT	4620
GTTGACATGC	GAACAAATCC	AGGTGATCCC	AGAGAAGCAC	AGGAACAGCA	AAATGACGGG	4680
AAAGGACGTG	GAAACAAGGC	AGCAAAACGA	GACCTTCCAC	CAGCAAAGAC	TCATCTCATC	4740
CAAGAGGATA	TTCTACCTTA	TTGTAGACCT	ACTTTTCCAA	CATCAAATAA	TCCCAGAGAT	4800
CCCAGTTCCT	CAAGCTCAAT	GTCATCAAGA	GGATCAGGAA	GCAGACAAAG	AGAACAAGCA	4860
AATGTAGGTC	GAAGAAATAT	TGCAGAAATG	CAGGTACTTG	GAGGATATGA	AAGAGGAGAA	4920
GATAATAATG	AAGAATTAGA	GGAAACTGAA	AGCTGA			4956

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1651 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met	Lys	Trp	Lys	His	Val	Pro	Phe	Leu	Val	Met	Ile	Ser	Leu	Leu	Ser
1				5				10					15		
Leu	Ser	Pro	Asn	His	Leu	Phe	Leu	Ala	Gln	Leu	Ile	Pro	Asp	Pro	Glu
			20					25					30		
Asp	Val	Glu	Arg	Gly	Asn	Asp	His	Gly	Thr	Pro	Ile	Pro	Thr	Ser	Asp
		35					40					45			
Asn	Asp	Asp	Asn	Ser	Leu	Gly	Tyr	Thr	Gly	Ser	Arg	Leu	Arg	Gln	Glu
		50				55				60					
Asp	Phe	Pro	Pro	Arg	Ile	Val	Glu	His	Pro	Ser	Asp	Leu	Ile	Val	Ser
65				70					75					80	
Lys	Gly	Glu	Pro	Ala	Thr	Leu	Asn	Cys	Lys	Ala	Glu	Gly	Arg	Pro	Thr
			85					90						95	
Pro	Thr	Ile	Glu	Trp	Tyr	Lys	Gly	Gly	Glu	Arg	Val	Glu	Thr	Asp	Lys

										100						105						110				
Asp	Asp	Pro	Arg	Ser	His	Arg	Met	Leu	Leu	Pro	Ser	Gly	Ser	Leu	Phe											
										115			120			125										
Phe	Leu	Arg	Ile	Val	His	Gly	Arg	Lys	Ser	Arg	Pro	Asp	Glu	Gly	Val											
										130			135			140										
Tyr	Val	Cys	Val	Ala	Arg	Asn	Tyr	Leu	Gly	Glu	Ala	Val	Ser	His	Asn											
145											150			155			160									
Ala	Ser	Leu	Glu	Val	Ala	Ile	Leu	Arg	Asp	Asp	Phe	Arg	Gln	Asn	Pro											
										165			170			175										
Ser	Asp	Val	Met	Val	Ala	Val	Gly	Glu	Pro	Ala	Val	Met	Glu	Cys	Gln											
										180			185			190										
Pro	Pro	Arg	Gly	His	Pro	Glu	Pro	Thr	Ile	Ser	Trp	Lys	Lys	Asp	Gly											
										195			200			205										
Ser	Pro	Leu	Asp	Asp	Lys	Asp	Glu	Arg	Ile	Thr	Ile	Arg	Gly	Gly	Lys											
										210			215			220										
Leu	Met	Ile	Thr	Tyr	Thr	Arg	Lys	Ser	Asp	Ala	Gly	Lys	Tyr	Val	Cys											
225											230			235			240									
Val	Gly	Thr	Asn	Met	Val	Gly	Glu	Arg	Glu	Ser	Glu	Val	Ala	Glu	Leu											
										245			250			255										
Thr	Val	Leu	Glu	Arg	Pro	Ser	Phe	Val	Lys	Arg	Pro	Ser	Asn	Leu	Ala											
										260			265			270										
Val	Thr	Val	Asp	Asp	Ser	Ala	Glu	Phe	Lys	Cys	Glu	Ala	Arg	Gly	Asp											
										275			280			285										
Pro	Val	Pro	Thr	Val	Arg	Trp	Arg	Lys	Asp	Asp	Gly	Glu	Leu	Pro	Lys											
										290			295			300										
Ser	Arg	Tyr	Glu	Ile	Arg	Asp	Asp	His	Thr	Leu	Lys	Ile	Arg	Lys	Val											
305											310			315			320									
Thr	Ala	Gly	Asp	Met	Gly	Ser	Tyr	Thr	Cys	Val	Ala	Glu	Asn	Met	Val											
										325			330			335										
Gly	Lys	Ala	Glu	Ala	Ser	Ala	Thr	Leu	Thr	Val	Gln	Glu	Pro	Pro	His											
										340			345			350										
Phe	Val	Val	Lys	Pro	Arg	Asp	Gln	Val	Val	Ala	Leu	Gly	Arg	Thr	Val											
										355			360			365										
Thr	Phe	Gln	Cys	Glu	Ala	Thr	Gly	Asn	Pro	Gln	Pro	Ala	Ile	Phe	Trp											
										370			375			380										
Arg	Arg	Glu	Gly	Ser	Gln	Asn	Leu	Leu	Phe	Ser	Tyr	Gln	Pro	Pro	Gln											
385											390			395			400									
Ser	Ser	Ser	Arg	Phe	Ser	Val	Ser	Gln	Thr	Gly	Asp	Leu	Thr	Ile	Thr											
										405			410			415										
Asn	Val	Gln	Arg	Ser	Asp	Val	Gly	Tyr	Tyr	Ile	Cys	Gln	Thr	Leu	Asn											
										420			425			430										
Val	Ala	Gly	Ser	Ile	Ile	Thr	Lys	Ala	Tyr	Leu	Glu	Val	Thr	Asp	Val											
										435			440			445										
Ile	Ala	Asp	Arg	Pro	Pro	Pro	Val	Ile	Arg	Gln	Gly	Pro	Val	Asn	Gln											
										450			455			460										
Thr	Val	Ala	Val	Asp	Gly	Thr	Phe	Val	Leu	Ser	Cys	Val	Ala	Thr	Gly											
465											470			475			480									
Ser	Pro	Val	Pro	Thr	Ile	Leu	Trp	Arg	Lys	Asp	Gly	Val	Leu	Val	Ser											
										485			490			495										
Thr	Gln	Asp	Ser	Arg	Ile	Lys	Gln	Leu	Glu	Asn	Gly	Val	Leu													

	530					535					540				
Phe	Gly	Val	Pro	Val	Gln	Pro	Pro	Arg	Pro	Thr	Asp	Pro	Asn	Leu	Ile
545					550					555					560
Pro	Ser	Ala	Pro	Ser	Lys	Pro	Glu	Val	Thr	Asp	Val	Ser	Arg	Asn	Thr
				565					570					575	
Val	Thr	Leu	Ser	Trp	Gln	Pro	Asn	Leu	Asn	Ser	Gly	Ala	Thr	Pro	Thr
			580					585					590		
Ser	Tyr	Ile	Ile	Glu	Ala	Phe	Ser	His	Ala	Ser	Gly	Ser	Ser	Trp	Gln
		595					600					605			
Thr	Val	Ala	Glu	Asn	Val	Lys	Thr	Glu	Thr	Ser	Ala	Ile	Lys	Gly	Leu
	610					615					620				
Lys	Pro	Asn	Ala	Ile	Tyr	Leu	Phe	Leu	Val	Arg	Ala	Ala	Asn	Ala	Tyr
625					630					635					640
Gly	Ile	Ser	Asp	Pro	Ser	Gln	Ile	Ser	Asp	Pro	Val	Lys	Thr	Gln	Asp
				645					650					655	
Val	Leu	Pro	Thr	Ser	Gln	Gly	Val	Asp	His	Lys	Gln	Val	Gln	Arg	Glu
			660					665					670		
Leu	Gly	Asn	Ala	Val	Leu	His	Leu	His	Asn	Pro	Thr	Val	Leu	Ser	Ser
		675					680					685			
Ser	Ser	Ile	Glu	Val	His	Trp	Thr	Val	Asp	Gln	Gln	Ser	Gln	Tyr	Ile
	690					695					700				
Gln	Gly	Tyr	Lys	Ile	Leu	Tyr	Arg	Pro	Ser	Gly	Ala	Asn	His	Gly	Glu
705					710					715					720
Ser	Asp	Trp	Leu	Val	Phe	Glu	Val	Arg	Thr	Pro	Ala	Lys	Asn	Ser	Val
				725					730					735	
Val	Ile	Pro	Asp	Leu	Arg	Lys	Gly	Val	Asn	Tyr	Glu	Ile	Lys	Ala	Arg
			740					745					750		
Pro	Phe	Phe	Asn	Glu	Phe	Gln	Gly	Ala	Asp	Ser	Glu	Ile	Lys	Phe	Ala
		755					760					765			
Lys	Thr	Leu	Glu	Glu	Ala	Pro	Ser	Ala	Pro	Pro	Gln	Gly	Val	Thr	Val
	770					775					780				
Ser	Lys	Asn	Asp	Gly	Asn	Gly	Thr	Ala	Ile	Leu	Val	Ser	Trp	Gln	Pro
785					790					795					800
Pro	Pro	Glu	Asp	Thr	Gln	Asn	Gly	Met	Val	Gln	Glu	Tyr	Lys	Val	Trp
				805					810					815	
Cys	Leu	Gly	Asn	Glu	Thr	Arg	Tyr	His	Ile	Asn	Lys	Thr	Val	Asp	Gly
			820					825					830		
Ser	Thr	Phe	Ser	Val	Val	Ile	Pro	Phe	Leu	Val	Pro	Gly	Ile	Arg	Tyr
		835					840					845			
Ser	Val	Glu	Val	Ala	Ala	Ser	Thr	Gly	Ala	Gly	Ser	Gly	Val	Lys	Ser
	850					855					860				
Glu	Pro	Gln	Phe	Ile	Gln	Leu	Asp	Ala	His	Gly	Asn	Pro	Val	Ser	Pro
865					870					875					880
Glu	Asp	Gln	Val	Ser	Leu	Ala	Gln	Gln	Ile	Ser	Asp	Val	Val	Lys	Gln
				885					890						

				965					970				975
Ala	Asp	Thr	Trp	Pro	Asn	Thr	Gly	Asn	Asn	His	Asn	Asp	Cys
				980					985				990
Ser	Cys	Cys	Thr	Ala	Gly	Asn	Gly	Asn	Ser	Asp	Ser	Asn	Leu
				995				1000				1005	
Tyr	Ser	Arg	Pro	Ala	Asp	Cys	Ile	Ala	Asn	Tyr	Asn	Asn	Gln
				1010			1015				1020		
Asn	Lys	Gln	Thr	Asn	Leu	Met	Leu	Pro	Glu	Ser	Thr	Val	Tyr
				1025		1030				1035			1040
Val	Asp	Leu	Ser	Asn	Lys	Ile	Asn	Glu	Met	Lys	Thr	Phe	Asn
				1045					1050				1055
Asn	Leu	Lys	Asp	Gly	Arg	Phe	Val	Asn	Pro	Ser	Gly	Gln	Pro
				1060				1065				1070	
Tyr	Ala	Thr	Thr	Gln	Leu	Ile	Gln	Ser	Asn	Leu	Ser	Asn	Asn
				1075			1080					1085	
Asn	Gly	Ser	Gly	Asp	Ser	Gly	Glu	Lys	His	Trp	Lys	Pro	Leu
				1090			1095				1100		
Gln	Lys	Gln	Glu	Val	Ala	Pro	Val	Gln	Tyr	Asn	Ile	Val	Glu
				1105		1110				1115			1120
Lys	Leu	Asn	Lys	Asp	Tyr	Arg	Ala	Asn	Asp	Thr	Val	Pro	Pro
				1125					1130				1135
Pro	Tyr	Asn	Gln	Ser	Tyr	Asp	Gln	Asn	Thr	Gly	Gly	Ser	Tyr
				1140				1145				1150	
Ser	Asp	Arg	Gly	Ser	Ser	Thr	Ser	Gly	Ser	Gln	Gly	His	Lys
				1155			1160				1165		
Ala	Arg	Thr	Pro	Lys	Val	Pro	Lys	Gln	Gly	Gly	Met	Asn	Trp
				1170			1175				1180		
Leu	Leu	Pro	Pro	Pro	Pro	Ala	His	Pro	Pro	Pro	His	Ser	Asn
				1185		1190				1195			1200
Glu	Tyr	Asn	Ile	Ser	Val	Asp	Glu	Ser	Tyr	Asp	Gln	Glu	Met
				1205					1210				1215
Pro	Val	Pro	Pro	Ala	Arg	Met	Tyr	Leu	Gln	Gln	Asp	Glu	Leu
				1220				1225				1230	
Glu	Glu	Asp	Glu	Arg	Gly	Pro	Thr	Pro	Pro	Val	Arg	Gly	Ala
				1235			1240					1245	
Ser	Pro	Ala	Ala	Val	Ser	Tyr	Ser	His	Gln	Ser	Thr	Ala	Thr
				1250			1255				1260		
Pro	Ser	Pro	Gln	Glu	Glu	Leu	Gln	Pro	Met	Leu	Gln	Asp	Cys
				1265		1270				1275			1280
Glu	Thr	Gly	His	Met	Gln	His	Gln	Pro	Asp	Arg	Arg	Arg	Gln
				1285					1290				1295
Ser	Pro	Pro	Pro	Pro	Pro	Arg	Pro	Ile	Ser	Pro	Pro	His	Thr
				1300				1305				1310	
Tyr	Ile	Ser	Gly	Pro	Leu	Val	Ser	Asp	Met	Asp	Thr	Asp	Ala
				1315			1320					1325	
Glu	Glu	Glu	Asp	Glu	Ala	Asp	Met	Glu	Val	Ala	Lys	Met	Gln
				1330			1335				1340		
Arg	Leu	Leu	Leu	Arg	Gly	Leu	Glu	Gln	Thr	Pro	Ala	Ser	Ser
				1345		1350				1355			1360
Asp	Leu	Glu	Ser	Ser	Val	Thr	Gly	Ser	Met	Ile	Asn	Gly	Trp
				1365					1370				1375
Ala	Ser	Glu	Glu	Asp	Asn	Ile	Ser	Ser	Gly	Arg	Ser	Ser	Val
				1380				1385				1390	
Ser	Asp	Gly	Ser	Phe	Phe	Thr	Asp	Ala	Asp	Phe	Ala	Gln	Ala

1395	1400	1405
Ala Ala Ala Glu Tyr Ala Gly Leu Lys Val Ala Arg Arg Gln Met Gln		
1410	1415	1420
Asp Ala Ala Gly Arg Arg His Phe His Ala Ser Gln Cys Pro Arg Pro		
1425	1430	1435
Thr Ser Pro Val Ser Thr Asp Ser Asn Met Ser Ala Ala Val Met Gln		
1445	1450	1455
Lys Thr Arg Pro Ala Lys Lys Leu Lys His Gln Pro Gly His Leu Arg		
1460	1465	1470
Arg Glu Thr Tyr Thr Asp Asp Leu Pro Pro Pro Pro Val Pro Pro Pro		
1475	1480	1485
Ala Ile Lys Ser Pro Thr Ala Gln Ser Lys Thr Gln Leu Glu Val Arg		
1490	1495	1500
Pro Val Val Val Pro Lys Leu Pro Ser Met Asp Ala Arg Thr Asp Arg		
1505	1510	1515
Ser Ser Asp Arg Lys Gly Ser Ser Tyr Lys Gly Arg Glu Val Leu Asp		
1525	1530	1535
Gly Arg Gln Val Val Asp Met Arg Thr Asn Pro Gly Asp Pro Arg Glu		
1540	1545	1550
Ala Gln Glu Gln Gln Asn Asp Gly Lys Gly Arg Gly Asn Lys Ala Ala		
1555	1560	1565
Lys Arg Asp Leu Pro Pro Ala Lys Thr His Leu Ile Gln Glu Asp Ile		
1570	1575	1580
Leu Pro Tyr Cys Arg Pro Thr Phe Pro Thr Ser Asn Asn Pro Arg Asp		
1585	1590	1595
Pro Ser Ser Ser Ser Ser Met Ser Ser Arg Gly Ser Gly Ser Arg Gln		
1605	1610	1615
Arg Glu Gln Ala Asn Val Gly Arg Arg Asn Ile Ala Glu Met Gln Val		
1620	1625	1630
Leu Gly Gly Tyr Glu Arg Gly Glu Asp Asn Asn Glu Glu Leu Glu Glu		
1635	1640	1645
Thr Glu Ser		
1650		

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1300 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 855..1187
- (D) OTHER INFORMATION: /note= "N signifies gap in sequence"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CAGATTGTTG CTCAAGGTCG AACAGTGACA TTTCCCTGTG AAATAAAGG AAACCCACAG	60
CCAGCTGTTT TTTGGCAGAA AGAAGGCAGC CAGAACCTAC TTTTCCCAA CCAACCCACAG	120
CAGCCCAACA GTAGATGCTC AGTGTACCA ACTGGAGACC TCACAATCAC CAACATTCAA	180
CGTTCCGACG CGGGTTACTA CATCTGCCAG GCTTTAACTG TGGCAGGAAG CATTTTAGCA	240
AAAGCTCAAC TGGAGGTTAC TGATGTTTTG ACAGATAGAC CTCCACCTAT AATTCTACAA	300
GGCCCAGCCA ACCAAACGCT GGCAGTGGAT GGTACAGCGT TACTGAAATG TAAAGCCACT	360
GGTGATCCTC TTCCTGTAAT TAGCTGGTTA AAGGAGGGAT TTACTTTTCC GGGTAGAGAT	420
CCAAGAGCAA CAATTCAAGA GCAAGGCACA CTGCAGATTA AGAATTTACG GATTTCTGAT	480

ACTGGCACTT ATACTTGTGT GGCTACAAGT TCAAGTGGAG AGGCTTCCTG GAGTGCAGTG 540
 CTGGATGTGA CAGAGTCTGG AGCAACAATC AGTAAAAACT ATGATTTAAG TGACCTGCCA 600
 GGGCCACCAT CCAAACCGCA AGTCACTGAT GTTACTAAGA ACAGTGTAC CTTGTCCTGG 660
 CAGCCAGGTA CCCCTGGAAC CCTTCCAGCA AGTGCATATA TCATTGAGGC TTTCAGCCAA 720
 TCAGTGAGCA ACAGCTGGCA GACCGTGGCA AACCATGTAA AGACCACCCT CTATACTGTA 780
 AGAGGACTGC GGCCCAATAC AATCTACTTA TTCATGGTCA GAGCGATCAA CCCCAGGTY 840
 TCAGTGACCC AAGTNAACC ACAGAAAAAC AATGGATCCA CTTGGGCCAA TGTCCCTCTA 900
 CCTCCCCCCC CAGTCCAGCC CCTTCCTGGC ACGGAGCTGG AACACTATGC AGTGGAACAA 960
 CAAGAAAATG GCTATGACAG TGATAGCTGG TGCCCACCAT TGCCAGTACA AACTTACTTA 1020
 CACCAAGGTC TGGAAGATGA ACTGGAAGAA GATGATGATA GGGTCCCAAC ACCTCCTGTT 1080
 CGAGGCGTGG CTTCTTCTCC TGCTATCTCC TTTGGACAGC AGTCCACTGC AACTCTTACT 1140
 CCATCCCCAC GGGGAAGAGAT GCAACCCATG CTGCAGGCTT CACCTNTTTA CCTCCTCTCA 1200
 AAGACCTCGA CCTACCAGCC CATTTTCTAC TGACAGTAAC ACCAGTGCAG CCCTGAGTCA 1260
 AAGTCAGAGG CCTCGGCCCA CTAAAAACA CAAGGGAGGG 1300

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 434 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site

- (B) LOCATION: 285..396

- (D) OTHER INFORMATION: /note= "Xaa signifies gap in sequence"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Gln Ile Val Ala Gln Gly Arg Thr Val Thr Phe Pro Cys Glu Thr Lys
 1 5 10 15
 Gly Asn Pro Gln Pro Ala Val Phe Trp Gln Lys Glu Gly Ser Gln Asn
 20 25 30
 Leu Leu Phe Pro Asn Gln Pro Gln Gln Pro Asn Ser Arg Cys Ser Val
 35 40 45
 Ser Pro Thr Gly Asp Leu Thr Ile Thr Asn Ile Gln Arg Ser Asp Ala
 50 55 60
 Gly Tyr Tyr Ile Cys Gln Ala Leu Thr Val Ala Gly Ser Ile Leu Ala
 65 70 75 80
 Lys Ala Gln Leu Glu Val Thr Asp Val Leu Thr Asp Arg Pro Pro Pro
 85 90 95
 Ile Ile Leu Gln Gly Pro Ala Asn Gln Thr Leu Ala Val Asp Gly Thr
 100 105 110
 Ala Leu Leu Lys Cys Lys Ala Thr Gly Asp Pro Leu Pro Val Ile Ser
 115 120 125
 Trp Leu Lys Glu Gly Phe Thr Phe Pro Gly Arg Asp Pro Arg Ala Thr
 130 135 140
 Ile Gln Glu Gln Gly Thr Leu Gln Ile Lys Asn Leu Arg Ile Ser Asp
 145 150 155 160
 Thr Gly Thr Tyr Thr Cys Val Ala Thr Ser Ser Ser Gly Glu Ala Ser
 165 170 175
 Trp Ser Ala Val Leu Asp Val Thr Glu Ser Gly Ala Thr Ile Ser Lys
 180 185 190
 Asn Tyr Asp Leu Ser Asp Leu Pro Gly Pro Pro Ser Lys Pro Gln Val
 195 200 205
 Thr Asp Val Thr Lys Asn Ser Val Thr Leu Ser Trp Gln Pro Gly Thr

210	215	220
Pro Gly Thr Leu Pro	Ala Ser Ala Tyr Ile	Ile Glu Ala Phe Ser Gln
225	230	235
Ser Val Ser Asn Ser	Trp Gln Thr Val Ala	Asn His Val Lys Thr Thr
245	250	255
Leu Tyr Thr Val Arg	Gly Leu Arg Pro Asn	Thr Ile Tyr Leu Phe Met
260	265	270
Val Arg Ala Ile Asn	Pro Lys Val Ser Val	Thr Gln Xaa Lys Pro Gln
275	280	285
Lys Asn Asn Gly Ser	Thr Trp Ala Asn Val	Pro Leu Pro Pro Pro Pro
290	295	300
Val Gln Pro Leu Pro	Gly Thr Glu Leu Glu	His Tyr Ala Val Glu Gln
305	310	315
Gln Glu Asn Gly Tyr	Asp Ser Asp Ser Trp	Cys Pro Pro Leu Pro Val
325	330	335
Gln Thr Tyr Leu His	Gln Gly Leu Glu Asp	Glu Leu Glu Glu Asp Asp
340	345	350
Asp Arg Val Pro Thr	Pro Pro Val Arg Gly	Val Ala Ser Ser Pro Ala
355	360	365
Ile Ser Phe Gly Gln	Gln Ser Thr Ala Thr	Leu Thr Pro Ser Pro Arg
370	375	380
Glu Glu Met Gln Pro	Met Leu Gln Ala Ser	Pro Xaa Phe Thr Ser Ser
385	390	395
Gln Arg Pro Arg Pro	Thr Ser Pro Phe Ser	Thr Asp Ser Asn Thr Ser
405	410	415
Ala Ala Leu Ser Gln	Ser Gln Arg Pro Arg	Pro Thr Lys Lys His Lys
420	425	430
Gly Gly		

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 444 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GCCCAGGCAG	TTGCTGCAGC	TGCGGAGTAT	GCGGGCCTGA	AAGTGGCTCG	CCGCCAAATG	60
CAAGATGCTG	CTGGCCGCCG	CCACTTCCAT	GCCTCTCAGT	GCCCAAGGCC	CACGAGTCCT	120
GTGTCCACAG	ACAGCAACAT	GAGTGCTGTT	GTGATCCAGA	AAGCCAGACC	CGCCAAGAAG	180
CAGAAACACC	AGCCAGGACA	TCTGCGCAGG	GAAGCCTACG	CAGATGATCT	TCCACCCCCT	240
CCAGTGCCAC	CACCTGCTAT	AAAATCGCCC	ACTGTCCAGT	CCAAGGCACA	GCTGGAGGTA	300
CGGCCTGTCA	TGGTGCCAAA	ACTCGCGTCT	ATAGAAGCAA	GGACAGATAG	ATCGTCAGAC	360
AGAAAAGGAG	GCAGTTACAA	GGGGAGAGAA	GCTCTGGATG	GAAGACAAGT	CACTGACCTG	420
CGAACAAATC	CAAGTGACCC	CAGA				444

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 148 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Ala	Gln	Ala	Val	Ala	Ala	Ala	Ala	Glu	Tyr	Ala	Gly	Leu	Lys	Val	Ala
1				5					10					15	
Arg	Arg	Gln	Met	Gln	Asp	Ala	Ala	Gly	Arg	Arg	His	Phe	His	Ala	Ser
			20					25					30		
Gln	Cys	Pro	Arg	Pro	Thr	Ser	Pro	Val	Ser	Thr	Asp	Ser	Asn	Met	Ser
		35					40					45			
Ala	Val	Val	Ile	Gln	Lys	Ala	Arg	Pro	Ala	Lys	Lys	Gln	Lys	His	Gln
	50					55					60				
Pro	Gly	His	Leu	Arg	Arg	Glu	Ala	Tyr	Ala	Asp	Asp	Leu	Pro	Pro	Pro
65					70				75					80	
Pro	Val	Pro	Pro	Pro	Ala	Ile	Lys	Ser	Pro	Thr	Val	Gln	Ser	Lys	Ala
				85				90					95		
Gln	Leu	Glu	Val	Arg	Pro	Val	Met	Val	Pro	Lys	Leu	Ala	Ser	Ile	Glu
			100					105					110		
Ala	Arg	Thr	Asp	Arg	Ser	Ser	Asp	Arg	Lys	Gly	Gly	Ser	Tyr	Lys	Gly
		115					120					125			
Arg	Glu	Ala	Leu	Asp	Gly	Arg	Gln	Val	Thr	Asp	Leu	Arg	Thr	Asn	Pro
	130					135					140				
Ser	Asp	Pro	Arg												
145															

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 909 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met	Phe	Asn	Arg	Lys	Thr	Leu	Leu	Cys	Thr	Ile	Leu	Leu	Val	Leu	Gln
1				5					10					15	
Ala	Val	Ile	Arg	Ser	Phe	Cys	Glu	Asp	Ala	Ser	Asn	Leu	Ala	Pro	Val
			20					25					30		
Ile	Ile	Glu	His	Pro	Ile	Asp	Val	Val	Val	Ser	Arg	Gly	Ser	Pro	Ala
		35				40						45			
Thr	Leu	Asn	Cys	Gly	Ala	Lys	Pro	Ser	Thr	Ala	Lys	Ile	Thr	Trp	Tyr
	50					55				60					
Lys	Asp	Gly	Gln	Pro	Val	Ile	Thr	Asn	Lys	Glu	Gln	Val	Asn	Ser	His
65					70				75					80	
Arg	Ile	Val	Leu	Asp	Thr	Gly	Ser	Leu	Phe	Leu	Leu	Lys	Val	Asn	Ser
			85					90					95		
Gly	Lys	Asn	Gly	Lys	Asp	Ser	Asp	Ala	Gly	Ala	Tyr	Tyr	Cys	Val	Ala
		100						105					110		
Ser	Asn	Glu	His	Gly	Glu	Val	Lys	Ser	Asn	Glu	Gly	Ser	Leu	Lys	Leu
	115						120					125			
Ala	Met	Leu	Arg	Glu	Asp	Phe	Arg	Val	Arg	Pro	Arg	Thr	Val	Gln	Ala
	130					135					140				
Leu	Gly	Gly	Glu	Met	Ala	Val	Leu	Glu	Cys	Ser	Pro	Pro	Arg	Gly	Phe
145					150				155					160	
Pro	Glu	Pro	Val	Val	Ser	Trp	Arg	Lys	Asp	Asp	Lys	Glu	Leu	Arg	Ile
			165					170					175		
Gln	Asp	Met	Pro	Arg	Tyr	Thr	Leu	His	Ser	Asp	Gly	Asn	Leu	Ile	Ile
		180						185					190		

Asp	Pro	Val	Asp	Arg	Ser	Asp	Ser	Gly	Thr	Tyr	Gln	Cys	Val	Ala	Asn		
		195					200					205					
Asn	Met	Val	Gly	Glu	Arg	Val	Ser	Asn	Pro	Ala	Arg	Leu	Ser	Val	Phe		
	210					215					220						
Glu	Lys	Pro	Lys	Phe	Glu	Gln	Glu	Pro	Lys	Asp	Met	Thr	Val	Asp	Val		
225					230					235					240		
Gly	Ala	Ala	Val	Leu	Phe	Asp	Cys	Arg	Val	Thr	Gly	Asp	Pro	Gln	Pro		
				245					250					255			
Gln	Ile	Thr	Trp	Lys	Arg	Lys	Asn	Glu	Pro	Met	Pro	Val	Thr	Arg	Ala		
			260					265					270				
Tyr	Ile	Ala	Lys	Asp	Asn	Arg	Gly	Leu	Arg	Ile	Glu	Arg	Val	Gln	Pro		
	275						280					285					
Ser	Asp	Glu	Gly	Glu	Tyr	Val	Cys	Tyr	Ala	Arg	Asn	Pro	Ala	Gly	Thr		
	290					295					300						
Leu	Glu	Ala	Ser	Ala	His	Leu	Arg	Val	Gln	Ala	Pro	Pro	Ser	Phe	Gln		
305					310					315					320		
Thr	Lys	Pro	Ala	Asp	Gln	Ser	Val	Pro	Ala	Gly	Gly	Thr	Ala	Thr	Phe		
				325					330					335			
Glu	Cys	Thr	Leu	Val	Gly	Gln	Pro	Ser	Pro	Ala	Tyr	Phe	Trp	Ser	Lys		
			340					345					350				
Glu	Gly	Gln	Asp	Leu	Leu	Phe	Pro	Ser	Tyr	Val	Ser	Ala	Asp	Gly			
	355					360					365						
Arg	Thr	Lys	Val	Ser	Pro	Thr	Gly	Thr	Leu	Thr	Ile	Glu	Glu	Val	Arg		
	370					375					380						
Gln	Val	Asp	Glu	Gly	Ala	Tyr	Val	Cys	Ala	Gly	Met	Asn	Ser	Ala	Gly		
385					390					395					400		
Ser	Ser	Leu	Ser	Lys	Ala	Ala	Leu	Lys	Val	Thr	Thr	Lys	Ala	Val	Thr		
				405					410					415			
Gly	Asn	Thr	Pro	Ala	Lys	Pro	Pro	Pro	Thr	Ile	Glu	His	Gly	His	Gln		
			420					425					430				
Asn	Gln	Thr	Leu	Met	Val	Gly	Ser	Ser	Ala	Ile	Leu	Pro	Cys	Gln	Ala		
	435					440						445					
Ser	Gly	Lys	Pro	Thr	Pro	Gly	Ile	Ser	Trp	Leu	Arg	Asp	Gly	Leu	Pro		
	450					455					460						
Ile	Asp	Ile	Thr	Asp	Ser	Arg	Ile	Ser	Gln	His	Ser	Thr	Gly	Ser	Leu		
465					470				475						480		
His	Ile	Ala	Asp	Leu	Lys	Lys	Pro	Asp	Thr	Gly	Val	Tyr	Thr	Cys	Ile		
				485					490					495			
Ala	Lys	Asn	Glu	Asp	Gly	Glu	Ser	Thr	Trp	Ser	Ala	Ser	Leu	Thr	Val		
			500					505					510				
Glu	Asp	His	Thr	Ser	Asn	Ala	Gln	Phe	Val	Arg	Met	Pro	Asp	Pro	Ser		
	515					520						525					
Asn	Phe	Pro	Ser	Ser	Pro	Thr	Gln	Pro	Ile	Ile	Val	Asn	Val	Thr	Asp		
	530					535					540						
Thr	Glu	Val	Glu	Leu	His	Trp	Asn	Ala	Pro	Ser	Thr	Ser	Gly	Ala	Gly		
545					550					555					560		
Pro	Ile	Thr	Gly	Tyr	Ile	Ile	Gln	Tyr	Tyr	Ser	Pro	Asp	Leu	Gly	Gln		
				565				570						575			
Thr	Trp	Phe	Asn	Ile	Pro	Asp	Tyr	Val	Ala	Ser	Thr	Glu	Tyr	Arg	Ile		
			580					585					590				
Lys	Gly	Leu	Lys	Pro	Ser	His	Ser	Tyr	Met	Phe	Val	Ile	Arg	Ala	Glu		
	595					600					605						
Asn	Glu	Lys	Gly	Ile	Gly	Thr	Pro	Ser	Val	Ser	Ser	Ala	Leu	Val	Thr		
	610					615					620						

Thr	Ser	Lys	Pro	Ala	Ala	Gln	Val	Ala	Leu	Ser	Asp	Lys	Asn	Lys	Met
625					630					635					640
Asp	Met	Ala	Ile	Ala	Glu	Lys	Arg	Leu	Thr	Ser	Glu	Gln	Leu	Ile	Lys
				645					650						655
Leu	Glu	Glu	Val	Lys	Thr	Ile	Asn	Ser	Thr	Ala	Val	Arg	Leu	Phe	Trp
			660					665					670		
Lys	Lys	Arg	Lys	Leu	Glu	Glu	Leu	Ile	Asp	Gly	Tyr	Tyr	Ile	Lys	Trp
		675					680					685			
Arg	Gly	Pro	Pro	Arg	Thr	Asn	Asp	Asn	Gln	Tyr	Val	Asn	Val	Thr	Ser
	690					695					700				
Pro	Ser	Thr	Glu	Asn	Tyr	Val	Val	Ser	Asn	Leu	Met	Pro	Phe	Thr	Asn
705					710					715					720
Tyr	Glu	Phe	Phe	Val	Ile	Pro	Tyr	His	Ser	Gly	Val	His	Ser	Ile	His
				725					730						735
Gly	Ala	Pro	Ser	Asn	Ser	Met	Asp	Val	Leu	Thr	Ala	Glu	Ala	Pro	Pro
			740					745					750		
Ser	Leu	Pro	Pro	Glu	Asp	Val	Arg	Ile	Arg	Met	Leu	Asn	Leu	Thr	Thr
		755					760					765			
Leu	Arg	Ile	Ser	Trp	Lys	Ala	Pro	Lys	Ala	Asp	Gly	Ile	Asn	Gly	Ile
	770					775					780				
Leu	Lys	Gly	Phe	Gln	Ile	Val	Ile	Val	Gly	Gln	Ala	Pro	Asn	Asn	Asn
785					790					795					800
Arg	Asn	Ile	Thr	Thr	Asn	Glu	Arg	Ala	Ala	Ser	Val	Thr	Leu	Phe	His
				805					810						815
Leu	Val	Thr	Gly	Met	Thr	Tyr	Lys	Ile	Arg	Val	Ala	Ala	Arg	Ser	Asn
			820					825					830		
Gly	Gly	Val	Gly	Val	Ser	His	Gly	Thr	Ser	Glu	Val	Ile	Met	Asn	Gln
		835					840					845			
Asp	Thr	Leu	Glu	Lys	His	Leu	Ala	Ala	Gln	Gln	Glu	Asn	Glu	Ser	Phe
	850					855					860				
Leu	Tyr	Gly	Leu	Ile	Asn	Lys	Ser	His	Val	Pro	Val	Ile	Val	Ile	Val
865					870					875					880
Ala	Ile	Leu	Ile	Ile	Phe	Val	Val	Ile	Ile	Ile	Ala	Tyr	Cys	Tyr	Trp
				885					890						895
Arg	Asn	Ser	Arg	Asn	Ser	Asp	Gly	Lys	Asp	Arg	Ser	Phe			
			900					905							